

Tuberculosis resistance acquisition in space and time: an analysis of globally diverse *M. tuberculosis* whole genome sequences

Yasha Ektefaie¹, Avika Dixit^{2,3}, Luca Freschi², Maha Farhat^{2,4}

¹University of California Berkeley, Berkeley CA, ²Harvard Medical School, Boston MA, ³Boston Children's Hospital, Boston MA, ⁴Massachusetts General Hospital, Boston, MA

Yasha Ektefaie

Affiliations: University of California Berkeley, Berkeley CA

Address: 120 Sproul Hall, Berkeley CA 94720, USA

Avika Dixit, MBBS MPH MBI

Affiliations: Boston Children's Hospital, Boston MA; Harvard Medical School, Boston MA

Address: 300 Longwood Ave, Mailstop 3103, Boston MA 02115, USA

Luca Freschi, PhD MSc

Affiliations: Harvard Medical School, Boston BA

Address: 10 Shattuck St, Boston MA 02115, USA

Maha R. Farhat, MD MSc

Affiliations: Harvard Medical School, Boston MA; Massachusetts General Hospital, Boston MA

Address: 10 Shattuck St, Boston MA 02115, USA

Corresponding Author:

Maha R. Farhat

10 Shattuck St, Boston MA 02130, USA

Email: maha_farhat@hms.harvard.edu

Research in context

Evidence before this study

Acquisition and spread of drug-resistance by *Mycobacterium tuberculosis* (MTB) varies across countries. Local factors driving evolution of drug resistance in MTB are not well studied.

Added value of this study

We applied molecular dating to 6,099 global MTB patient isolates and found the order of resistance acquisition to be consistent across the countries examined, *i.e.* acquisition of isoniazid resistance first followed by rifampicin and streptomycin followed by resistance to other drugs. In all countries with data available there was evidence for transmission of resistant strains from patient-to-patient and in the majority for extended periods of time (>20 years). Countries with lower gross wealth indices were found to have more recent resistance acquisition to the drug rifampicin. Based on the resistance patterns identified in our study we estimate that commercial diagnostic tests vary considerably in sensitivity for second-line resistance diagnosis by country.

Implications of all available evidence

The longevity of resistant MTB in many parts of the world emphasizes its fitness for transmission and its continued threat to public health. The association between country wealth and recent resistance acquisition emphasizes the need for continued investment in TB care delivery and surveillance programs. Geographically relevant diagnostics that take into account a country's unique distribution of resistance are necessary.

Abstract

Background: *Mycobacterium tuberculosis* (MTB) whole genome sequencing data can provide insights into temporal and geographic trends in resistance acquisition and inform public health interventions.

Methods: We curated a set of clinical MTB isolates with high quality sequencing and culture-based drug susceptibility data spanning four lineages and more than 20 countries. We constructed geographic and lineage specific MTB phylogenies and used Bayesian molecular dating to infer the most-recent-common-susceptible-ancestor age for 4,869 instances of resistance to 10 drugs.

Findings: Of 8,550 isolates curated, 6,099 from 15 countries met criteria for molecular dating. The number of independent resistance acquisition events was lower than the number of resistant isolates across all countries, suggesting ongoing transmission of drug resistance. Ancestral age distributions supported the presence of old resistance, ≥ 20 years prior, in the majority of countries. A consistent order of resistance acquisition was observed globally starting with resistance to isoniazid, but resistance ancestral age varied by country. We found a direct correlation between country wealth and resistance age ($R^2 = 0.47$, $P\text{-value} = 0.014$). Amplification of fluoroquinolone and second-line injectable resistance among multidrug-resistant isolates is estimated to have occurred very recently (median ancestral age 4.7 years IQR 1.9-9.8 prior to sample collection). We found the sensitivity of commercial molecular diagnostics for second-line resistance to vary significantly by country ($P\text{-value} < 0.0003$).

Interpretation: Our results highlight that both resistance transmission and amplification are contributing to disease burden globally but are variable by country. The observation that wealthier nations are more likely to have old resistance suggests that programmatic improvements can reduce resistance amplification, but that fit resistant strains can circulate for decades subsequently.

Funding: This work was supported by the NIH BD2K grant K01 ES026835, a Harvard Institute of Global Health Burke Fellowship (MF), Boston Children's Hospital OFD/BTREC/CTREC Faculty Career Development Fellowship and Bushrod H. Campbell and Adah F. Hall Charity Fund/Charles A. King Trust Postdoctoral Fellowship (AD).

Keywords: tuberculosis, drug resistance, whole genome sequencing

Introduction

Tuberculosis (TB) defines a global epidemic that takes more lives than any other infection due to a single pathogen¹. The emergence of multidrug-resistant (MDR) and extensively drug-resistant (XDR) resistant TB presents a major hurdle to efforts in accelerating TB decline. Halting the transmission of drug-resistant (DR) TB has been a major focus of studies addressing this hurdle². But the epidemic is ultimately defined by local factors that remain understudied in many parts of the world³. The study of geographic and temporal heterogeneity of the DR-TB epidemic can provide insights into these local factors as key drivers of MDR-TB prevalence and persistence in the community, including programmatic and bacterial factors. This understanding is key to future disease control and prevention of antibiotic resistance development.

Over the past decade, increased uptake of molecular and whole genome sequencing (WGS) technologies, and their application to *Mycobacterium tuberculosis* (MTB) clinical isolates has offered novel insights into pathogen biology and diversity in the context of human infection⁴⁻⁷. The application of WGS has allowed us to better understand the genetic determinants of drug resistance (DR) within MTB⁸. The detection of these genetic determinants using molecular technologies that include WGS is now increasingly adopted for TB resistance diagnosis in many parts of the world⁹ and is beginning to replace the more biohazardous and time consuming culture based drug susceptibility tests (DST). The study of isolates sampled from epidemiological outbreaks or from the same host over time has allowed the estimation of MTB's molecular clock rate, or temporal rate of accumulation of fixed genome-level variation^{10,11}. The application of this rate to new WGS data from isolates collected for surveillance has helped improve transmission inference and molecular dating of specific evolutionary events such as resistance acquisition or lineage divergence^{10,12-13}.

We sought to use a large clinical collection of MTB WGS and resistance phenotype data to study how, when, and where resistance was acquired on a global scale. Using a Bayesian implementation of coalescent theory, we estimate and compare dates of resistance acquisition for MDR/XDR isolates across 15 different countries. We use the recency of resistance acquisition as a measure of fitness of the circulating strains in their respective environments and study the effect of country wealth, as a proxy for TB control programme funding, on the recency of resistance acquisition at a macro level. We also assess the distribution of unexplained MTB phenotypic resistance across 20 countries, to evaluate the accuracy and geographic heterogeneity of molecular detection of common MTB genetic resistance determinants, and discuss implications for DR-TB control.

Methods

Further details available in the supplementary material.

Data and quality control

We compiled a 10,299 MTB WGS dataset with culture based DST (phenotypic) data using public databases (Patric¹⁴, ReSeqTB¹⁵) and literature curation^{11–13,16–26}. A summary table with the phenotypic data is available online at <https://github.com/farhat-lab/resdata>.

Genomic analysis/variant calling

We used a previously validated genomic analysis pipeline for MTB described by Ezewudo *et al.*²⁷ with modifications as detailed in the supplement.

Drug resistance definitions

Drugs were labelled as follows: isoniazid (INH), ethambutol (EMB), rifamycins (rifampicin or rifabutin) (RIF), streptomycin (STR), pyrazinamide (PZA), fluoroquinolones (FLQ) (includes moxifloxacin, ciprofloxacin, ofloxacin), second-line injectables (SLIs) (includes kanamycin, amikacin, capreomycin), ethionamide/prothionamide (ETH), and cycloserine (CYS). Para-aminosalicylic acid was not analysed due to the paucity of data. Isolates not tested for susceptibility to both INH and RIF were excluded from the assessment of DR frequency by country and lineage. Isolates resistant to both INH and RIF were labelled MDR. Those resistant to INH, RIF, FLQ and SLIs were labelled XDR.

Estimating resistance acquisition dates

Isolates were separated into 179 groups corresponding to a single drug, lineage and source country, referred to hereafter as a 'group'. Genetic diversity was computed as the average pairwise genetic distance within a group. To accurately date resistance acquisition, a drug-geography-lineage group was analysed only if it consisted of at least 10 isolates, $\geq 20\%$ of isolates were susceptible and ≥ 1 isolate was resistant. To exclude isolates that only represent outbreak settings and didn't carry more long-term information about resistance, we excluded groups with a genetic diversity score < 1 standard deviation from the mean genetic diversity score measured across all groups. Supplementary methods detail the phylogeny construction and the estimation of the age of the most recent *susceptible* common ancestor (MRSCA) in years prior to isolation of the clinical sample(s).

Distribution of Resistance Mutations

We compared the expected sensitivity and specificity of mutations captured by commercial diagnostics (summarized from the literature in Table S9) and those based on more extensive lists of mutations in DR genes that can be captured using targeted or whole genome sequencing. We used three mutations lists for the latter (1) a set of 267 common resistance-associated mutations that we previously determined using randomForests²⁸ designated "RF-select WGS test", (2) a set of mutations determined using direct association²⁹ designated "DA-select WGS test", and (3) any non-synonymous mutation or noncoding mutation in known DR regions (Table S10) in a "all WGS test". We excluded previously described neutral/lineage associated mutations⁹.

Code

All code used in the analysis is publically accessible at https://github.com/farhat-lab/geo_dist_tb.

Results

Data and global lineage distribution

Of the 10,299 MTB clinical isolates with WGS and culture-based DST data available, 9,385 passed sequence quality criteria and of these 8,550 had country of origin data (Figure 1). The four major MTB lineages, 1-4, were well represented. A relatively high proportion, 42%, of United Kingdom (UK) isolates (n= 1873) belonged to Lineages-1 & 3 (Figure 2A). Overall, the non-Europe-America-Africa Lineages-1,3, and 2, comprised 40% of European isolates (n=3956) and 7% of North and South American isolates (n=1297).

Phenotypic resistance distribution

Of the 8,550 isolates, 568 isolates lacked either INH or RIF DST data. Out of the remaining 7,909 isolates, 5,022 were pan-susceptible, 2887 were resistant to one or more drugs (DR) and of these 1937 were resistant to INH and RIF (MDR) and 288 were MDR and resistant to an SLI and a FLQ (XDR). The 8,550 isolates originated from 52 countries. Of these, 23 countries were represented by >10 isolates with resistance data, 21/23 were found to have MDR isolates and 9/21 had XDR (Figure 2B). We compared the MDR frequency in our WGS based sample with the WHO reported MDR/RIF resistance (RR) rates for the latest year available³⁰. Out of the 21 countries, the confidence interval for the MDR-TB proportion in our sample overlapped with that of the WHO in 4 (19%) countries, was higher in 14 (67%) and lower in 3 (14%) of countries (Table S4). MDR rates by lineage were 3% for Lineage-1 (n=439), 48% for Lineage-2 (n=1085), 4% for Lineage-3 (n=760) and 23% for Lineage-4 (n=3358).

Molecular dating of Resistance Acquisition

Of the 8,550 isolates, 2,451 isolates appeared in groups that did not meet our dating requirements (Methods). The remaining 6,009, included 1,547 isolates resistant to one or more drugs and were grouped into 179 country/lineage/drug combinations. We estimated 4,869 MRSCA dates for 10 drugs across these 179 groups. The number of independent resistance acquisition events *i.e.* unique MRSCA dates, was consistently lower than the total number of dated resistance isolates suggesting ongoing transmission of drug resistant isolates (Table S11). We estimated a lower bound on the burden of resistance due to transmission ranging by country from ≥14% to ≥52% pooled across drugs (Methods, Table S11). The proportion of INH or RIF resistance attributed to transmission was the highest among the 10 drugs at ≥43% and ≥46% respectively pooled across countries (pooled from Table S11).

We examined the relative order of phenotypic resistance acquisition on a global scale. For INH, we found that resistance to INH on average developed before resistance to other drugs (Figure 3A-B). Median MRSCA for INH was 11.4 years prior to isolation (IQR 6.3-16.2) vs. 7.6 years (IQR 3.0 – 16.0) for RIF, Wilcoxon P-value <10⁻¹⁴. Median MRSCA ages for RIF and STR resistance (7.6 years, IQR 3.0 – 16.0 and 7.7 years, IQR 3.4 – 13.0 respectively) were second oldest and not statistically significant from each other (Wilcoxon P-value 0.31). The dating supported that EMB resistance followed the acquisition of RIF (Wilcoxon P-value < 10⁻⁶) at a median MRSCA age of 5.0 years prior (IQR 2.1 – 12.5), and that this was followed by resistance to PZA, ETH, FLQ, SLIs, or CYS (Figure 3A), amongst which MRSCA ages did not significantly differ (Figure 3B). We found no significant correlation between the median MRSCA dates and

the drug's date of introduction into clinical use with $R^2 = 0.04$ (F-test with 1 DF P-value= 0.60, Table S12).

We assessed the frequency of recent resistance amplification to PZA, EMB, FLQs and SLIs among MDR, *i.e.* to pre-XDR/XDR, within five years of sample isolation. Among the 11 countries with both MDR and pre-XDR/XDR isolates, we identified four countries (Peru, Russia, Sierra Leone, South Africa) with recent resistance amplification to PZA and EMB (>1% of MDR). The rates of recent amplification ranged from 2% (95% CI 1% - 4%) for PZA in Russia to 33% (95% CI 26% - 41%) for EMB in South Africa (Figure 5). Peru, Romania and South Africa were also measured to have recent resistance amplification to FLQs and SLIs (Figure 5). The median MRSCA age for FLQ or SLI resistance acquisition among MDR isolates was 4.7 years (IQR 1.9-9.8) prior to sample collection.

We found RIF to have the highest proportion of old resistance (MRSCA >20 years prior to isolation) at 17%, 197/1184 out of the total dated RIF resistance acquisition events. Old resistance was well distributed geographically and for RIF occurred in 9 of 12 countries with available dating data (Figure 4). Old FLQ resistance constituted 8% (24/311) of total dated isolates and spanned 6 of the 7 countries with available data.

We compared the geographic distribution of MRSCA ages restricting to four key drug classes, namely INH, RIF, SLIs and FLQs, and the five countries with the largest number of resistant isolates (Figure 4). MRSCA ages did not differ between the UK and China across all four drug classes. These two countries had the oldest median MRSCA across the five countries and four drug classes except for INH. MRSCA ages in the UK were a median of 13.0 years (IQR 10.7-19.0) for RIF, a median of 10.6 years (IQR 7.3-11.2) for SLIs and a median of 8.4 years (IQR 6.7-18.4) for FLQs (Figure 4). South Africa most consistently had the youngest median MRSCA for the four drug classes, but its MRSCA distribution was not significantly different from that of Peru (for FLQs and SLIs) and Russia (for SLIs) (Table S3). A similar geographic/age pattern was observed for the drugs PZA and EMB across these five countries (Table S3).

We examined if the geographic resistance age differences correlated with resources available for TB control programs using the gross domestic product per capita as a proxy. We found GDP to correlate significantly with an older RIF MRSCA date with $R^2 = 0.47$ (F-test with 1 DF P-value= 0.014) (Figure 6 & Table S8).

Distribution of Resistance Mutations:

We assessed the frequency of 267 resistance mutations previously determined to be important for resistance prediction²⁸ and their geographic distribution among the 8,550 isolates with country of origin and WGS data meeting quality criteria (Figure 1). Resistance mutation prevalence varied significantly by country. The most frequent INH causing mutation³¹, *katG* S315T was more frequent among phenotypically INH resistant isolates by DST (pheno-R) in Russia (84%, n=526) than in Peru (67%, n=760) (Fisher P-value 1×10^{-12}). The second most common INH resistance mutation -15 C>T *fabG1/inhA* promoter was more prevalent among INH pheno-R Peruvian isolates (20%) than in Russian isolates (8%) (Fisher P-value 7×10^{-9}). Twenty four of the 267 resistance mutations (9%) varied geographically to a larger extent than the mutation *fabG/inhA* promoter -15C>T (standard deviation 11%, frequency range 0-39%, Table S13). The mutation I491F was recently described to be common in Estawini³² and is not detectable by line-probe or

GeneXpert commercial molecular diagnostics. In our sample that did not contain data from Estawini, we calculated a standard deviation of 1% for the global frequency of I491F (range 0% - 4%) among RIF pheno-R isolates.

We calculated the proportion of pheno-R isolates that can be captured by the Hain Line-probe or GeneXpert commercial molecular diagnostics due to the presence of one or more mutations in their pooled target regions for the drugs INH, RIF, SLIs, and FLQ (Tables 1 and S14). Sensitivity was highest for RIF (90% of 2624) and lowest for FLQs (51% of 854). Specificity was consistently high (lowest for SLIs at 86%) (Table 1). Second-line sensitivity of commercial diagnostics differed significantly across countries (Table S14). FLQ sensitivity in Peru was 38% (n= 121) and 77% in South Africa (n=111) (Fisher P-value 1×10^{-6}). A similarly low sensitivity for SLI resistance was seen in Peru compared with South Africa (Fisher P-value 3×10^{-4}) (Table S14).

We examined if expanding the resistance mutation list to variants previously characterized in diverse global MTB genomic datasets using direct association²⁹ or random forests²⁸ can improve sensitivity and specificity in a “select WGS test.” The select WGS test improved sensitivity slightly for INH and SLIs with relatively preserved specificity (Table 1). In addition select WGS test allowed for prediction of resistance to other drugs not tested by commercial diagnostics: PZA, EMB and Streptomycin. For comparison, we assessed if including any non-silent variant in the resistance regions (excluding a select number of known lineage markers) was indeed inferior to the more informed ‘select WGS test’ reported previously. We found that this “all WGS test” only modestly improved sensitivity and at the expense of a larger decrease in specificity (Figure 7).

Discussion

Using 8,550 clinical MTB sequences with culture-based DST, we examined geographic trends in the DR-TB epidemic. Geographically, MTB lineages 1-4 were each represented in the continents of Europe, Asia and Africa providing evidence of disease spread across borders, likely driven by human migration³⁰. We found MDR-TB in nearly every country represented by more than 10 isolates. XDR isolates were found in half of these countries and spanning all five major continents. Lineage-2 had the highest percentage of MDR isolates in our sample followed by Lineages-4, 3 and 1. Using this diverse sample we dated more than 4869 resistance phenotypes across 4 lineages and 15 countries.

We found a consistent order of resistance acquisition globally among drug classes. The development of INH resistance was previously found to be a sentinel event heralding the development of MDR³³. Our results corroborate these findings using phenotypic resistance data and across a larger geographically diverse sample. After INH, we find that MTB acquires resistance to RIF/STR then EMB followed by PZA, ETH, FLQ, SLIs, or CYS. We found no correlation between the age of resistance acquisition and the year of clinical introduction of the drug but there may be multiple other causes for the observed order of resistance acquisition. Differences in mutation rates across drug targets or resistance genes have been postulated but shown to be an unlikely explanation for INH resistance arising first^{34,35}. Pharmacokinetic difference may result in higher risk for under-dosing³⁶ for some drugs and earlier resistance acquisition. Bacterial fitness costs are also variable across resistance mutations. For INH

resistance, mutations like *katG* S315T carry a low fitness cost and likely contribute to resistance arising earliest for this drug^{33,37,38}. The order of drug administration can explain dating differences between first-line (INH, RIF, EMB, PZA) and second (ETH, FLQ, SLIs) or third-line (CYS) resistance, as second-line drugs are usually only administered after resistance to first-line drugs is ascertained. Acquisition of resistance to INH first then RIF may also relate to their use for treatment of latent TB infection, leading to more exposure and selection pressure overall. However, because adoption of INH preventative therapy for latent TB remains low in many parts of the world, we expect it to be a lesser contributor to INH and RIF resistance rates³⁹. Lastly, the observation of contemporaneous acquisition of RIF and STR resistance is likely best explained by the effects of Category II TB treatment initially recommended in 1991⁴⁰. Category II is no longer recommended by the WHO but consists of adding streptomycin to the first-line drug regimen after treatment failure. Our dating supports that streptomycin resistance amplified among patients failing due to recent RIF resistance and/or MDR acquisition.

Published evidence supports that most resistant cases of MTB result from recent resistance acquisition in the host or are related to transmission⁴¹. Reactivation of resistant MTB disease acquired remotely (>2 years prior) is much less likely⁴². Thus the identification of isolates with old resistance suggests high fitness for continued transmission between human hosts. Most countries with available data had isolates with resistance dated more than 20 years prior. This is also supported by our phylogenetic assessment where we estimate a lower bound of TB resistance due to transmission to range between 14-52% across countries with available data. As our approach cannot distinguish between resistance importation through human migration after transmission outside of the country and new resistance acquisition, these figures are underestimates of the true resistance burden due to transmission. Mathematical models of TB rates have previously predicted transmission to be a major driver of observed resistance rates⁴³, we present here WGS based evidence of the high burden of resistance transmission. Mathematical models have also emphasized that drug resistant strain fitness is a key parameter that dictates how the resistance epidemic will unfold. Our results support that >14-52% of isolates are fit and successfully transmitting patient-to-patient and in most countries there have been uninterrupted chains of resistance transmission for >20 years. These data emphasize the need to contain resistance transmission through improved diagnosis, treatment and other preventative strategies such as infection control and vaccine development.

In addition to transmission, we find evidence for recent resistance amplification, especially to second-line drugs mediating the transition from MDR to XDR-TB. XDR has considerably worse treatment outcome than susceptible TB and incurs more than 25 times the cost⁴⁴. We estimate that half of FLQ and SLI resistant isolates had acquired resistance within 4.7 years of isolation despite the promotion of directly-observed-therapy (DOT) by the WHO since 1994. As most FLQ and SLI resistant isolates are also MDR, our results also emphasize the need for better regimens to treat MDR that can prevent resistance amplification. By country, we found a significant correlation between the estimated age of resistance acquisition and per capita GDP, with more affluent countries having older ages of resistance. This unexpected correlation is likely driven by a combination of factors but the routine use of DST and close patient monitoring in the more well-resourced health systems are likely important contributors. Specifically, we found the UK and China to have the oldest resistance ages across the drugs. The Chinese national TB program budget was increased from \$98 million in 2002 to \$272 million in 2007⁴⁵

and a new policy for free TB diagnostics tests and drug use was introduced in 2004^{58,46}. This increased investment can explain the observed low rates of recent resistance acquisition in China^{30,47}.

Likely due to geographic differences in MTB lineage, transmission and resistance acquisition rates, we find 10% of assayed resistance mutations to have high geographic variance. We also found commercial diagnostics to vary in sensitivity for second-line drugs. Given recent reports about the accuracy of WGS for confirming susceptibility of MTB²⁹, we measured improvements in resistance sensitivity offered by including mutations outside of regions targeted by commercial diagnostics through direct association. This offered modest improvements in sensitivity with little to no change in specificity. We found a considerable number of indeterminate mutations in resistance regions, that when included with simple direct association improve sensitivity but at the expense of loss of specificity. The study of these variants through statistical models will likely further inform their diagnostic use in the future^{28,48}.

Our study has several limitations including the oversampling of DR isolates as evidenced by our comparison with WHO reported MDR rates. We tried to control for this by dating only in countries with at least 20% susceptible isolates and limiting dating of low diversity samples that represent unique outbreak settings and lack long term information about resistance acquisition. This may have resulted in underestimation of rates of recent resistance acquisition but despite this we were able to document recent resistance acquisition in many countries. Molecular dating is also reliant on the accurate estimation of the phylogenetic tree of MTB isolates and the molecular clock assumption. We thus used a rigorous approach to phylogenetic estimation and dating despite its computational and time cost⁴⁹. Our analysis also assumes the accuracy of culture based phenotypic DST, even though test to test variability is known to exist. We justify this as our data was curated from ReseqTB¹⁵ and studies where phenotypic testing was performed in national or supranational laboratories with rigorous quality control.

Overall, our results support that DR rates are fuelled by both recent resistance acquisition and ongoing transmission, and suggest the need for better detection, treatment and health system investment. In the future, reassessment of these patterns will be enabled by the sharing of systematically collected isolate data, data that is increasingly generated as by-products of TB surveillance and resistance diagnosis²⁹.

Figures and Tables

Figure 1: Flow diagram showing process of identification and exclusion of genomic data included in the study. WGS: Whole genome sequencing, QC: Quality control.

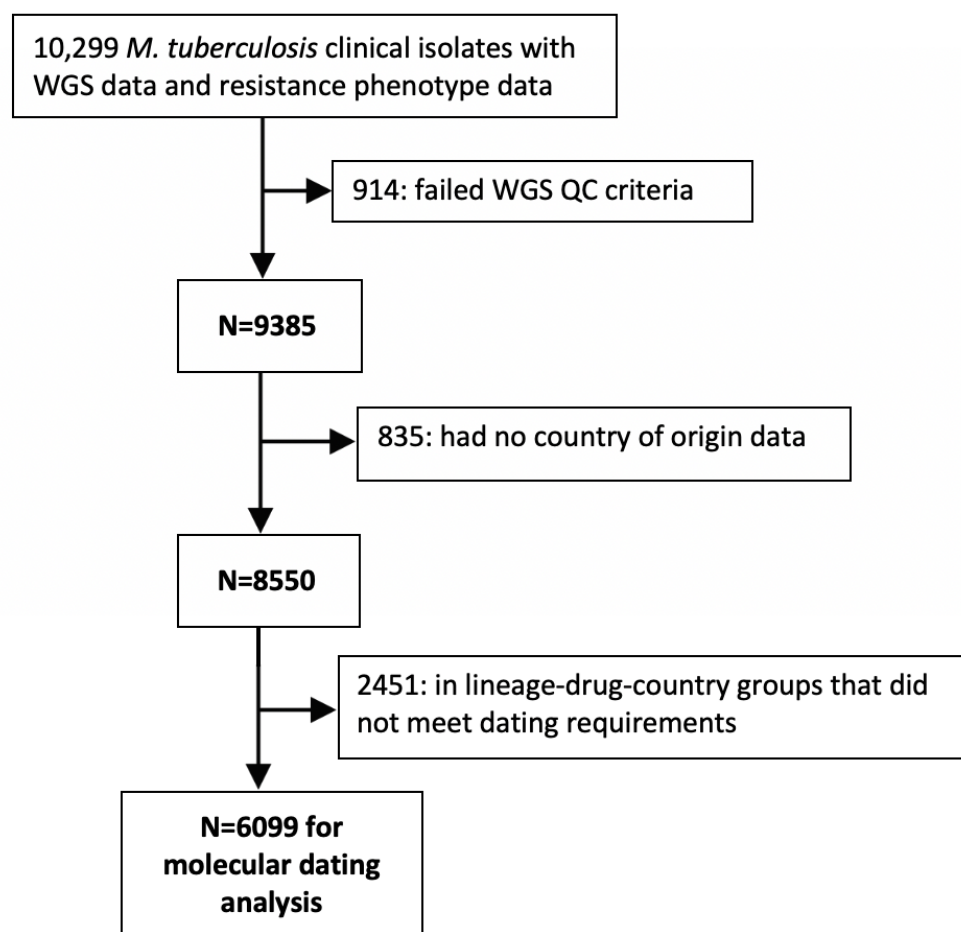
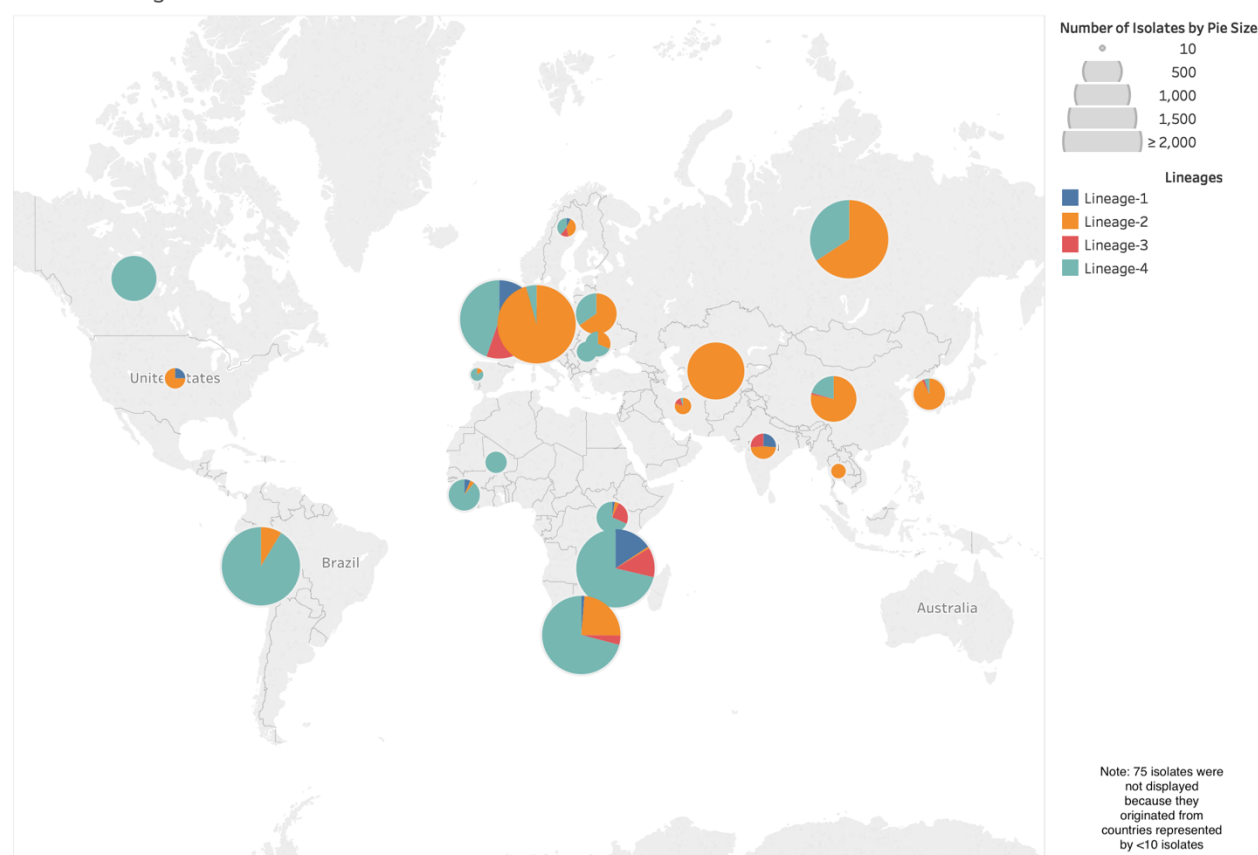


Figure 2: Global Distribution of *M. tuberculosis* in the study sample. Counts from countries represented by fewer than 10 isolates (n=75) not shown. **A:** Lineage distribution (n=8477). Pie charts represent the proportion of each lineage among isolates available from each country. Size of the pie is proportional to the number of isolates from each country. Detailed counts are in Table S1. **B:** Drug resistance distribution (n=7834). Pie charts provide the distribution of resistance patterns (S: Susceptible, MDR: Multidrug-resistant, XDR: Extensively drug resistant, INH Mono: mono resistant to isoniazid, STR Mono: mono resistant to streptomycin, Other R: resistance other than defined categories) by country. 75 isolates originated Pie size is proportional to the number of isolates in each country (Table S2).

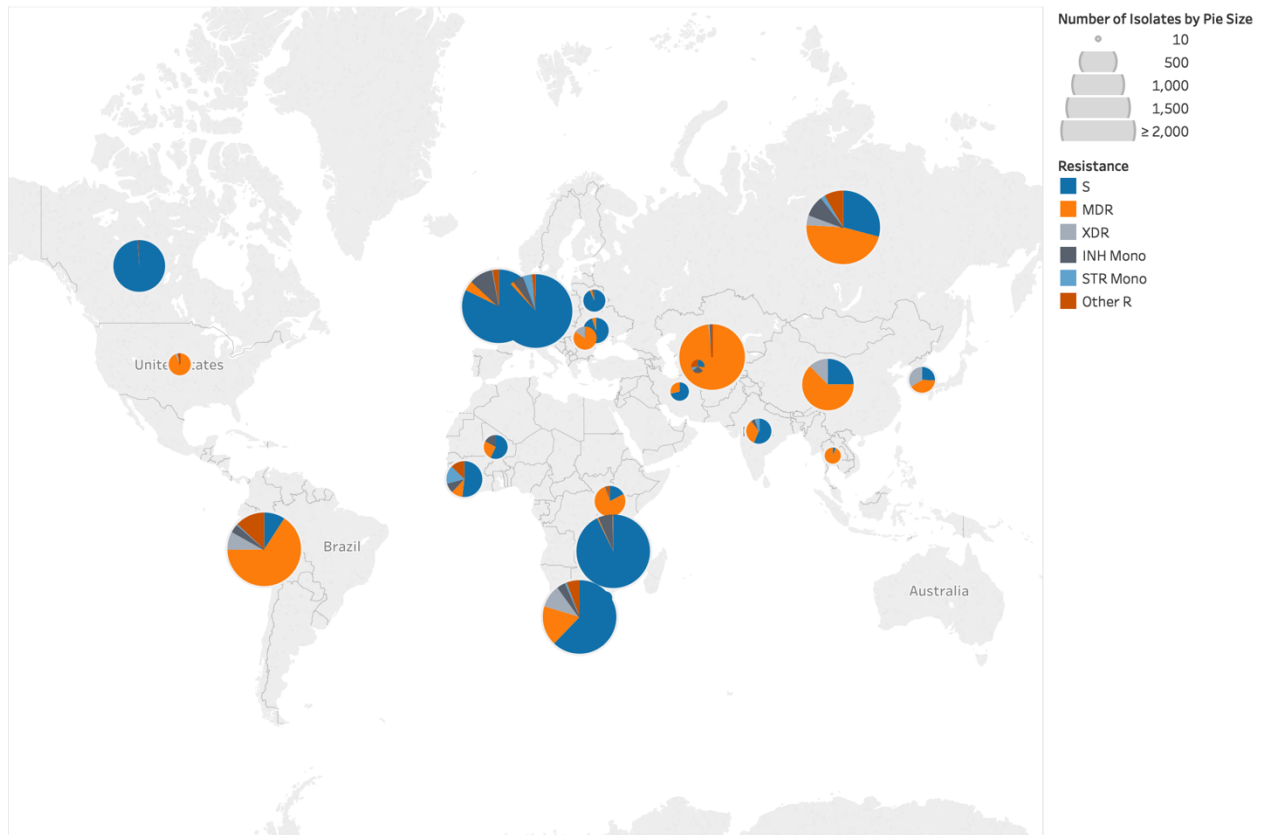
2A:

Global Lineage Distribution



345 2B:

Global Resistance Distribution



346

347

348

Figure 3A: MRSCA distribution by drug (n=4844). Boxplots showing range of MRSCA distribution globally for nine anti-tubercular drugs (MRSCA: Most recent susceptible common ancestor, INH: isoniazid, RIF: rifampicin, EMB: ethambutol, PZA: pyrazinamide, STR: streptomycin, FLQ: fluoroquinolones, SLIs: second-line injectables, ETH: ethionamide, CYS: cycloserine)

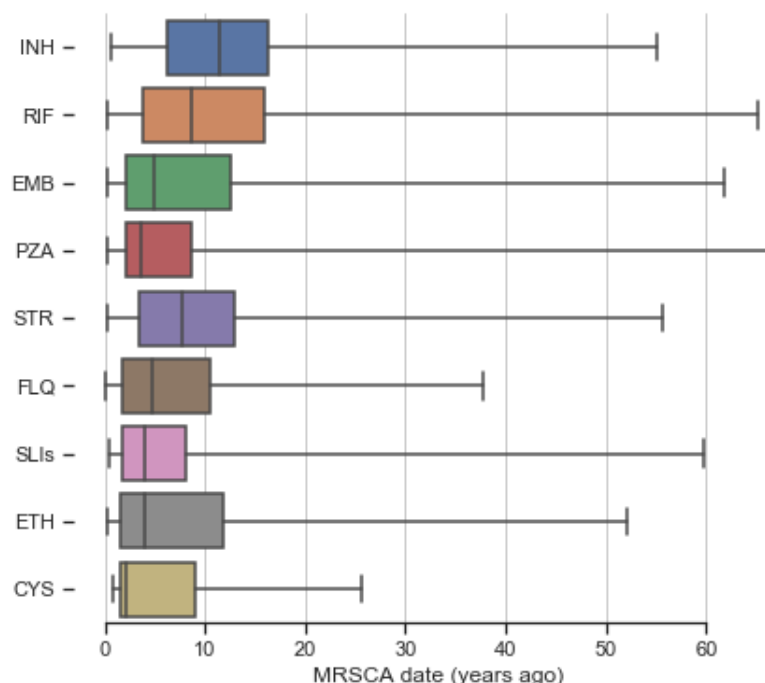


Figure 3B: Pairwise Wilcoxon rank sum tests comparing MRSCA ages by drug category. Dark red indicates P-value <0.001 (Bonferroni threshold); pink indicates P-value <0.01; white indicates P-value ≥0.01.

	INH	RIF	EMB	PZA	STR	FLQ	SLIs	ETH	CYS
INH		5 E-16	2 E-29	3 E-46	1 E-17	1 E-31	1 E-59	1 E-23	5 E-6
RIF			2 E-6	1 E-15	0.3	4 E-11	9 E-21	2 E-8	9 E-4
EMB				1 E-3	5 E-4	0.01	1 E-4	0.02	0.03
PZA					5 E-12	0.9	0.8	0.9	0.08
STR						5 E-8	2 E-16	6 E-6	3 E-3
FLQ							0.4	0.9	0.2
SLIs								0.5	0.1
ETH									0.2
CYS									

Figure 4: Box and whisker plots summarizing the MRSCA distribution per country. (Blue vertical line indicates year when drug was introduced (Table S5))

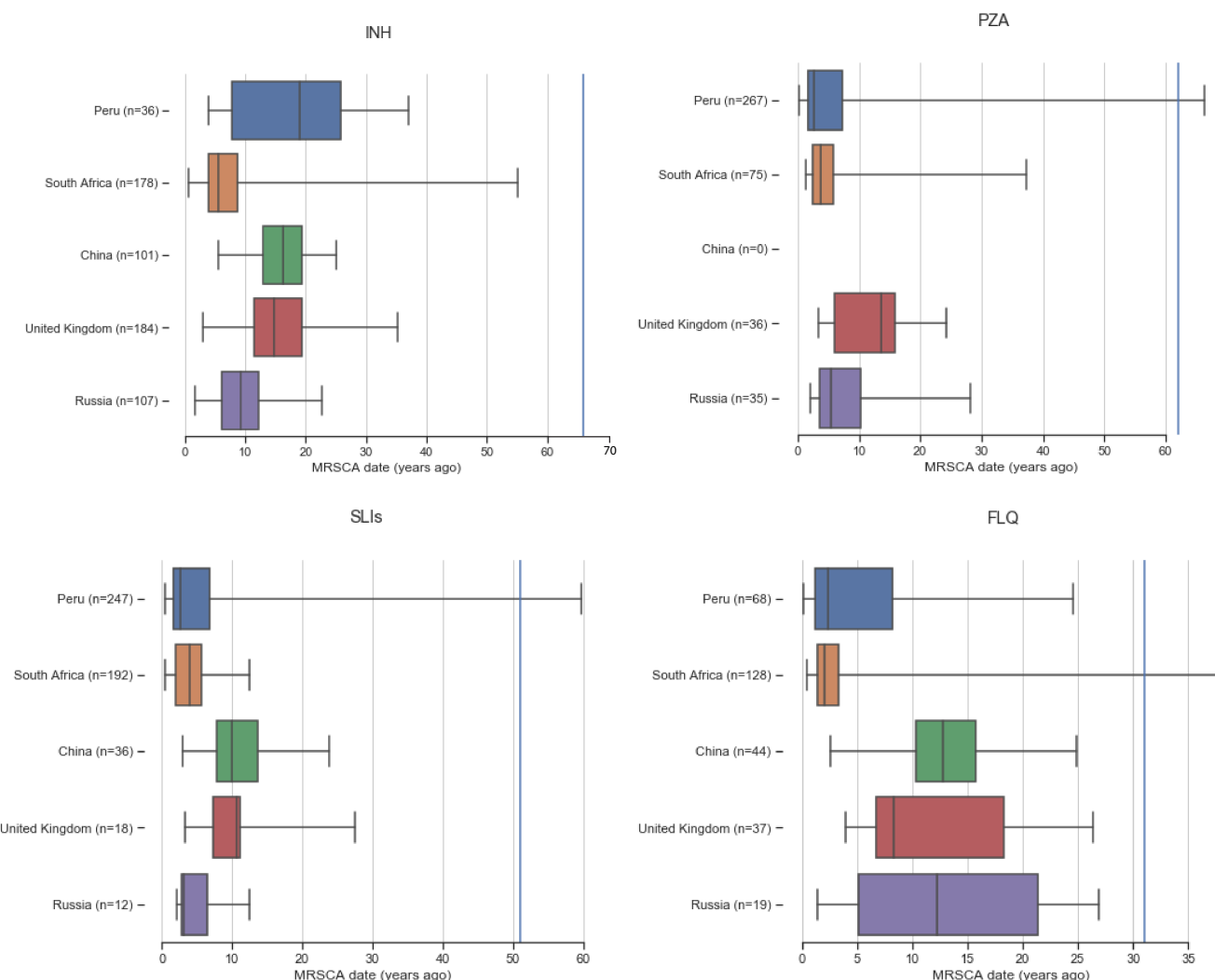


Figure 5: Proportion of MDR isolates with recent amplification of resistance to ethambutol (EMB), pyrazinamide (PZA), fluoroquinolones (FLQ) or second-line injectables (SLIs) by country (MRSCA age estimate <5 years ago). The legend lists the number of multidrug-resistant isolates analysed from each country. Error bars indicate 95% confidence intervals. Full data given in Supplementary Table 3. Four countries displayed a measurable proportion of recent FLQ and SLI amplification (95% CI does not include 0) – South Africa, Peru, Romania and China.

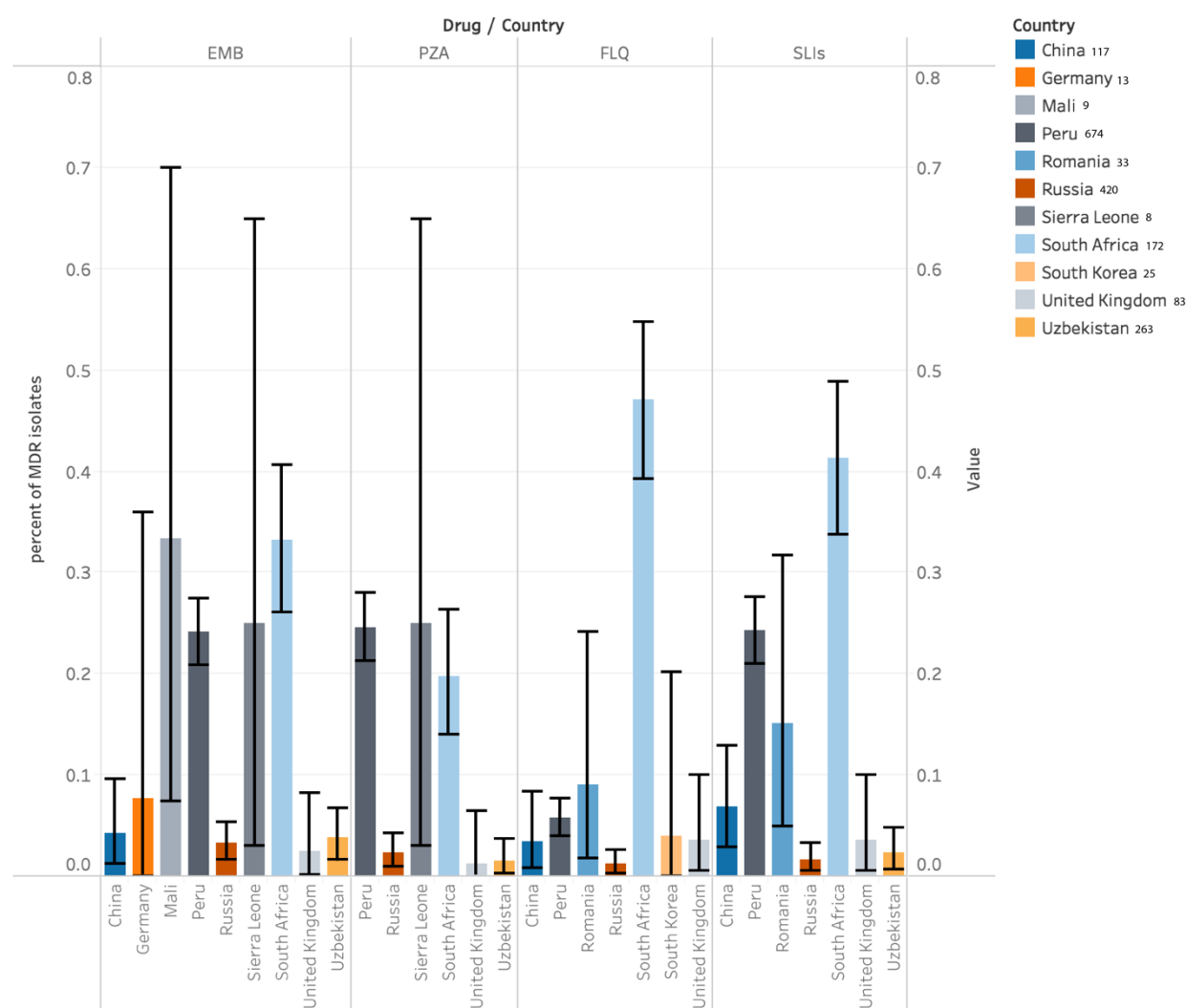


Figure 6: Median Rifamycin (RIF) most-recent-common-susceptible-ancestor (MRSCA) date vs Gross Domestic Product (GDP) per capita for 12 countries. Data plotted is provided in Supplementary Table 7 and includes other drugs than RIF. ($R^2 = 0.47$, F-test P-value (1 DF) = 0.014).

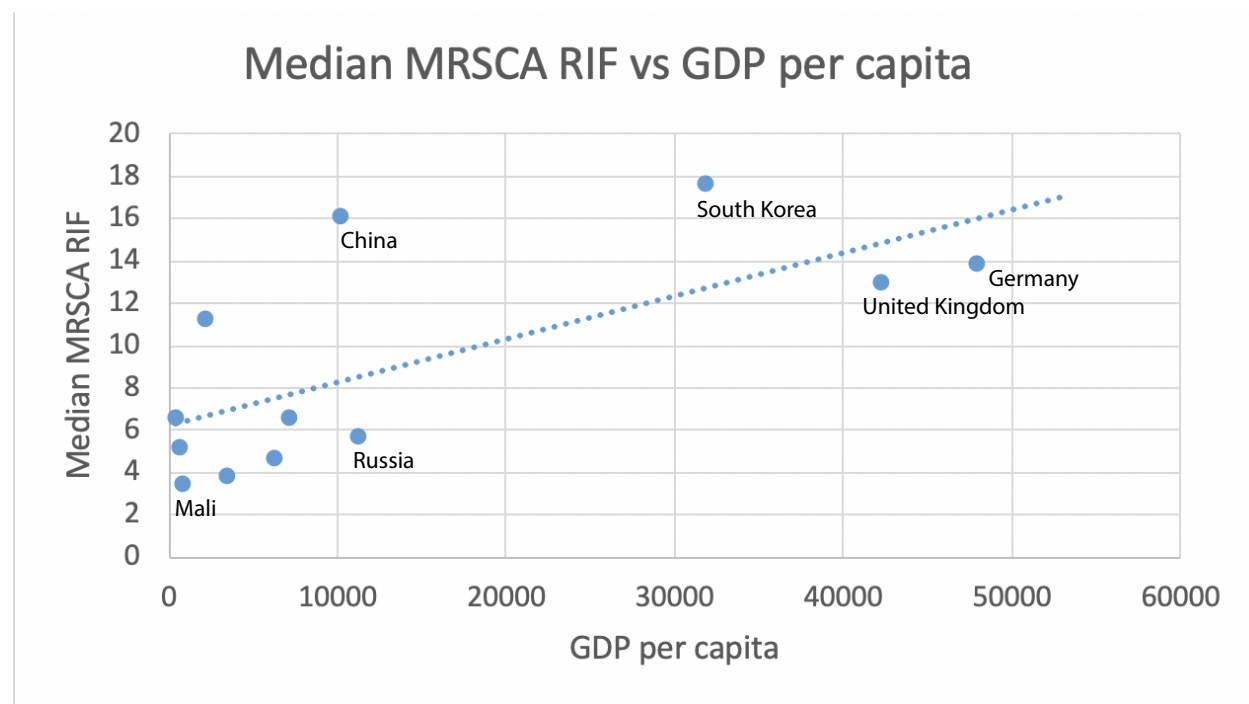


Table 1: Sensitivity and Specificity of commercial and WGS based tests for resistance diagnosis. Abbreviations defined in “drug resistance definitions” section of methods.

Drug	commercial test		RF-select WGS test		DA-select WGS test		all WGS test*	
	Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity
INH	83% (2759/3306)	93% (4834/5201)	88% (2900/3306)	92% (4780/5201)	89% (2956/3306)	92% (4776/5201)	92% (3029/3306)	64% (3314/5201)
RIF	90% (2354/2624)	93% (5361/5786)	91% (2385/2624)	92% (5341/5786)	91% (2395/2624)	92% (5338/5786)	92% (2405/2624)	89% (5178/5786)
FLQ	53% (452/854)	94% (2626/2790)	51% (439/854)	95% (2637/2790)	51% (440/854)	95% (2641/2790)	57% (488/854)	86% (2406/2790)
SLI	56% (517/921)	86% (2179/2547)	58% (535/921)	84% (2136/2547)	57% (524/921)	86% (2185/2547)	64% (594/921)	80% (2027/2547)
PZA			65% (862/1324)	95% (4660/4907)	75% (996/1324)	91% (4485/4907)	78% (1030/1324)	89% (4384/4907)
EMB			79% (1476/1863)	86% (4702/5441)	75% (1389/1863)	88% (4772/5441)	85% (1589/1863)	74% (4042/5441)
STR			76% (1601/2112)	86% (3204/3726)	77% (1627/2112)	87% (3223/3726)	91% (1928/2112)	71% (2657/3726)
Legend								
*	classifying FLQ and SLI as susceptible if no INH and no RIF resistance mutation found							
Sensitivity	Percent of resistant isolates classified as resistant							
Specificity	Percent of susceptible isolates classified as susceptible							

392 **Authors' contributions:**

393 Yasha Ektefaie conducted the data analysis, drafted and revised the manuscript.

394 All authors provided key edits to the manuscript.

395 Additionally:

396 Luca Freschi contributed to the data analysis.

397 Avika Dixit contributed to the data analysis.

398 Maha Farhat conceptualized the study, supervised the data analysis, reviewed, wrote and
399 edited the manuscript.

400

401 **Declaration of interests:**

402

403 Yasha Ektefaie: None

404 Luca Freschi: None

405 Avika Dixit: None

406 Maha Farhat: None

References:

1. World Health Organization. *Global Tuberculosis Report 2016*. (World Health Organization, 2016).
2. Churchyard, G. *et al.* What We Know About Tuberculosis Transmission: An Overview. *J Infect Dis* **216**, S629–S635 (2017).
3. Mathema, B. *et al.* Drivers of Tuberculosis Transmission. *J Infect Dis* **216**, S644–S653 (2017).
4. Farhat, M. R. *et al.* Genomic analysis identifies targets of convergent positive selection in drug-resistant *Mycobacterium tuberculosis*. *Nat. Genet.* (2013) doi:10.1038/ng.2747.
5. Farhat, M. R. *et al.* Genome wide association with quantitative resistance phenotypes in *Mycobacterium tuberculosis* reveals novel resistance genes and regulatory regions. *bioRxiv* 429159 (2018) doi:10.1101/429159.
6. Coll, F. *et al.* Genome-wide analysis of multi- and extensively drug-resistant *Mycobacterium tuberculosis*. *Nature Genetics* **50**, 307–316 (2018).
7. Nebenzahl-Guimaraes, H. *et al.* Transmissible *Mycobacterium tuberculosis* Strains Share Genetic Markers and Immune Phenotypes. *Am J Respir Crit Care Med* **195**, 1519–1527 (2016).
8. The CRyPTIC Consortium and the 100, 000 Genomes Project. Prediction of Susceptibility to First-Line Tuberculosis Drugs by DNA Sequencing. *New England Journal of Medicine* (2018) doi:10.1056/NEJMoa1800474.
9. Walker, T. M. *et al.* Whole-genome sequencing for prediction of *Mycobacterium tuberculosis* drug susceptibility and resistance: a retrospective cohort study. *Lancet Infect Dis* (2015) doi:10.1016/S1473-3099(15)00062-6.

10. Dixit, A. *et al.* Whole genome sequencing identifies bacterial factors affecting transmission of multidrug-resistant tuberculosis in a high-prevalence setting. *Scientific Reports* **9**, 5602 (2019).
11. Walker, T. M. *et al.* Whole-genome sequencing to delineate *Mycobacterium tuberculosis* outbreaks: a retrospective observational study. *Lancet Infect Dis* **13**, 137–146 (2013).
12. Gardy, J. L. *et al.* Whole-Genome Sequencing and Social-Network Analysis of a Tuberculosis Outbreak. *New England Journal of Medicine* **364**, 730–739 (2011).
13. Merker, M. *et al.* Evolutionary history and global spread of the *Mycobacterium tuberculosis* Beijing lineage. *Nature Genetics* **47**, 242–249 (2015).
14. Wattam, A. R. *et al.* Improvements to PATRIC, the all-bacterial Bioinformatics Database and Analysis Resource Center. *Nucleic Acids Res.* **45**, D535–D542 (2017).
15. Ezewudo, M. *et al.* Integrating standardized whole genome sequence analysis with a global *Mycobacterium tuberculosis* antibiotic resistance knowledgebase. *Sci Rep* **8**, 15382 (2018).
16. Chatterjee, A., Nilgiriwala, K., Saranath, D., Rodrigues, C. & Mistry, N. Whole genome sequencing of clinical strains of *Mycobacterium tuberculosis* from Mumbai, India: A potential tool for determining drug-resistance and strain lineage. *Tuberculosis* **107**, 63–72 (2017).
17. Manson, A. L. *et al.* *Mycobacterium tuberculosis* Whole Genome Sequences From Southern India Suggest Novel Resistance Mechanisms and the Need for Region-Specific Diagnostics. *Clin. Infect. Dis.* **64**, 1494–1501 (2017).
18. Walker, T. M. *et al.* Whole-genome sequencing for prediction of *Mycobacterium tuberculosis* drug susceptibility and resistance: a retrospective cohort study. *Lancet Infect. Dis.* **15**, 1193–1202 (2015).

19. Bryant, J. M. *et al.* Inferring patient to patient transmission of *Mycobacterium tuberculosis* from whole genome sequencing data. *BMC Infect. Dis.* **13**, 110 (2013).
20. Casali, N. *et al.* Evolution and transmission of drug-resistant tuberculosis in a Russian population. *Nat. Genet.* **46**, 279–286 (2014).
21. Clark, T. G. *et al.* Elucidating emergence and transmission of multidrug-resistant tuberculosis in treatment experienced patients by whole genome sequencing. *PLoS One* **8**, e83012 (2013).
22. Out-of-Africa migration and Neolithic coexpansion of *Mycobacterium tuberculosis* with modern humans | Nature Genetics. <https://www.nature.com/articles/ng.2744#s1>.
23. Perdigão, J. *et al.* Unraveling *Mycobacterium tuberculosis* genomic diversity and evolution in Lisbon, Portugal, a highly drug resistant setting. *BMC Genomics* **15**, 991 (2014).
24. Blouin, Y. *et al.* Significance of the identification in the Horn of Africa of an exceptionally deep branching *Mycobacterium tuberculosis* clade. *PLoS One* **7**, e52841 (2012).
25. Cohen, K. A. Evolution of Extensively Drug-Resistant Tuberculosis over Four Decades: Whole Genome Sequencing and Dating Analysis of *Mycobacterium tuberculosis* Isolates from KwaZulu-Natal. <https://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1001880>.
26. Zhang, H. *et al.* Genome sequencing of 161 *Mycobacterium tuberculosis* isolates from China identifies genes and intergenic regions associated with drug resistance. *Nat. Genet.* **45**, 1255–1260 (2013).
27. Ezewudo, M. *et al.* Integrating standardized whole genome sequence analysis with a global *Mycobacterium tuberculosis* antibiotic resistance knowledgebase. *Scientific Reports* **8**, 15382 (2018).

28. Farhat, M. R. *et al.* Genetic Determinants of Drug Resistance in *Mycobacterium tuberculosis* and Their Diagnostic Value. *Am J Respir Crit Care Med* **194**, 621–630 (2016).
29. Prediction of Susceptibility to First-Line Tuberculosis Drugs by DNA Sequencing. *New England Journal of Medicine* **379**, 1403–1415 (2018).
30. WHO | Tuberculosis country profiles. *WHO* <http://www.who.int/tb/country/data/profiles/en/>.
31. Nebenzahl-Guimaraes, H., Jacobson, K. R., Farhat, M. R. & Murray, M. B. Systematic review of allelic exchange experiments aimed at identifying mutations that confer drug resistance in *Mycobacterium tuberculosis*. *J Antimicrob Chemother* **69**, 331–342 (2014).
32. André, E. *et al.* Novel rapid PCR for the detection of Ile491Phe *rpoB* mutation of *Mycobacterium tuberculosis*, a rifampicin-resistance-conferring mutation undetected by commercial assays. *Clin. Microbiol. Infect.* **23**, 267.e5-267.e7 (2017).
33. Manson, A. L. *et al.* Genomic analysis of globally diverse *Mycobacterium tuberculosis* strains provides insights into the emergence and spread of multidrug resistance. *Nature Genetics* **49**, 395–402 (2017).
34. Bergval, I. L., Schuitema, A. R. J., Klatser, P. R. & Anthony, R. M. Resistant mutants of *Mycobacterium tuberculosis* selected in vitro do not reflect the in vivo mechanism of isoniazid resistance. *J. Antimicrob. Chemother.* **64**, 515–523 (2009).
35. Ford, C. B. *et al.* *Mycobacterium tuberculosis* mutation rate estimates from different lineages predict substantial differences in the emergence of drug-resistant tuberculosis. *Nat. Genet.* **45**, 784–790 (2013).
36. Jutte, P. C., Rutgers, S. R., Van Altena, R., Uges, D. R. & Van Horn, J. R. Penetration of isoniazid, rifampicin and pyrazinamide in tuberculous pleural effusion and psoas abscess. *Int. J. Tuberc. Lung Dis.* **8**, 1368–1372 (2004).

37. Ordway, D. J., Sonnenberg, M. G., Donahue, S. A., Belisle, J. T. & Orme, I. M. Drug-resistant strains of Mycobacterium tuberculosis exhibit a range of virulence for mice. *Infect Immun* **63**, 741–743 (1995).
38. Pym, A. S., Saint-Joanis, B. & Cole, S. T. Effect of katG mutations on the virulence of Mycobacterium tuberculosis and the implication for transmission in humans. *Infect. Immun.* **70**, 4955–4960 (2002).
39. Alsdurf, H., Hill, P. C., Matteelli, A., Getahun, H. & Menzies, D. The cascade of care in diagnosis and treatment of latent tuberculosis infection: a systematic review and meta-analysis. *Lancet Infect Dis* **16**, 1269–1278 (2016).
40. WHO | Eliminating the category II retreatment regimen from national tuberculosis programme guidelines: the Georgian experience. *WHO* <https://www.who.int/bulletin/volumes/90/1/11-092320/en/>.
41. Kendall, E. A., Fofana, M. O. & Dowdy, D. W. Burden of transmitted multidrug resistance in epidemics of tuberculosis: a transmission modelling analysis. *Lancet Respir Med* **3**, 963–972 (2015).
42. Behr, M. A., Edelstein, P. H. & Ramakrishnan, L. Revisiting the timetable of tuberculosis. *BMJ* **362**, k2738 (2018).
43. Kendall, E. A., Fojo, A. T. & Dowdy, D. W. Expected effects of adopting a 9 month regimen for multidrug-resistant tuberculosis: a population modelling analysis. *Lancet Respir Med* **5**, 191–199 (2017).
44. Marks, S. M. *et al.* Treatment Practices, Outcomes, and Costs of Multidrug-Resistant and Extensively Drug-Resistant Tuberculosis, United States, 2005–2007 - Volume 20, Number 5—May 2014 - Emerging Infectious Diseases journal - CDC. doi:10.3201/eid2005.131037.

- 522 45. WHO | Global tuberculosis reports (from 1997). *WHO*
523 http://www.who.int/tb/publications/global_report/archive/en/.
- 524 46. Health reform and development report. [http://www.gov.cn/ztl/2005-](http://www.gov.cn/ztl/2005-10/20/content_80720.htm)
525 [10/20/content_80720.htm](http://www.gov.cn/ztl/2005-10/20/content_80720.htm).
- 526 47. Zhao, Y. *et al.* National survey of drug-resistant tuberculosis in China. *N. Engl. J. Med.* **366**,
527 2161–2170 (2012).
- 528 48. Chen, M. L. *et al.* Beyond multidrug resistance: Leveraging rare variants with machine and
529 statistical learning models in Mycobacterium tuberculosis resistance prediction.
530 *EBioMedicine* **43**, 356–369 (2019).
- 531 49. Menardo, F., Duchêne, S., Brites, D. & Gagneux, S. The molecular clock of Mycobacterium
532 tuberculosis. *PLoS Pathog.* **15**, e1008067 (2019).
- 533

Supplementary Methods:

Data curation and quality control:

The isolate metadata including their geographic locations were downloaded using metatools ncbi (https://github.com/farhat-lab/metatools_ncbi). We also performed literature curation to fill the gaps in the NCBI geographic location data. The resulting table of the geographic locations of the isolates is available in Supplementary File 1. We excluded isolates that did not meet WGS quality control criteria as detailed below, had no geographic information or were not tested for phenotypic resistance to one or more drugs.

Genomic analysis/variant calling:

Briefly, reads were trimmed using PRINSEQ¹ setting average phred score threshold to 20. Raw read data was confirmed to belong to MTB complex using Kraken². Isolates with <90% mapping were excluded. Reads were aligned to H37Rv (GenBank NC000962.3) reference genome using BWA MEM³. Duplicate reads were removed using PICARD⁴. We excluded any isolates with coverage <95% of known drug resistance regions (*katG*, *inhA* & its promoter, *rpoB*, *embA*, *embB*, *embC* & *embB* promoter, *ethA*, *gyrA* and *gyrB*, *rrs*, *rpsL*, *gid*, *pncA*, *rpsA*, *eis* promoter) at 10x or higher. Variants were called using Pilon⁵ that uses local assembly to increase indel (insertions and deletions) call accuracy. This deviates from Ezewudo *et al.* that uses Samtools for variant calling⁶. The reference allele was implied if allele frequency was <75% or the Pilon filter was not PASS. Low confidence coordinates were filtered from all strains if >95% of strains did not have coverage of (trimmed reads) at least 10x at that site. Isolate lineage belonging to one of the seven main MTB lineages was confirmed using the Coll *et al.* SNP barcode⁷.

Drug resistance definitions and comparison with WHO reported resistance proportions:

The ‘Mono’ resistant designation was given to isolates that were resistant to only one specific drug and susceptible to all others that were tested, with the exception of the INH-mono resistant category that encompassed any isolates that were resistant to INH and/or STR but not to others that were tested. The ‘Other-R’ category was reserved for isolates that were resistant to some drugs but were neither INH or STR mono resistant, nor MDR or XDR. Isolates were labelled susceptible if they were susceptible to all drugs tested.

To compute exact binomial confidence intervals for MDR proportions by country we used the python library statsmodel⁸. To assess overlap with World Health Organization (WHO) estimates we determined whether the confidence intervals of our proportion intersected with that of the WHO. We labelled our estimate as high if our confidence interval was higher than the WHO, low if it was lower, and the same if they intersected.

Estimating resistance acquisition dates and lower bounds of resistance transmission:

A maximum likelihood tree was generated for each group via RAxML 8.2.11⁹ with H37Rv (NC000962.3) as the outgroup, starting from a neighbour-joining seed tree and assuming a generalized time reversible (GTR) nucleotide substitution model with the Γ distribution used to model site rate heterogeneity¹⁰. We bootstrapped the maximum likelihood tree 1000 times. The maximum likelihood tree was dated using BEAST v1.10.4¹¹ assuming a relaxed molecular clock with a log normal distribution and a mean rate of 0.5 SNP per genome per year based on prior published data¹². Sumtrees.py from the DendroPy library¹³ was then used to combine the output from the bootstrap analysis and that of BEAST to get our final dated phylogenetic tree with nodal bootstrap support.

We dated the most recent common ancestor between all the resistant isolates and their most closely related susceptible isolate. Accordingly, the dates of resistance acquisition will be referred to as the estimated age of the most recent *susceptible* common ancestor (MRSCA) in years prior to isolation of the clinical sample(s) throughout the text. We excluded resistant isolates with MRSCAs inferred at nodes with less than 50 bootstrap support.

We calculated the number of phylogenetically inferred resistance acquisition events (N_{aq}) per country and lineage as the number of unique MRSCAs identified. This was compared with the total resistant isolates that could be dated (N_{td}). Phylogenetically inferred unique resistance acquisition events for a particular country may be related to either *in host* evolution of new resistance or due to human migration/importation from another country with the latter still possibly related to transmission elsewhere. Thus, the following quantity represents a lower bound on the burden of resistance due to transmission for a particular country:

$$(N_{td} - N_{aq})/N_{td}$$

To estimate the order of resistance acquisition for different drugs we pooled the MRSCA dates by drug across countries and lineages. We compared the medians of the MRSCA distributions and performed pairwise Wilcoxon Rank Sum tests to assess for statistical significance, correcting for multiple testing using the Bonferroni approach. Interquartile ranges were calculated using the python package numpy¹⁴.

We correlated the median MRSCA date per drug pooled across countries against the date of drug introduction using linear regression as implemented in Microsoft Excel version 16.25.

Distribution of resistance mutations

We measured resistance mutations' geographic variance by calculating the proportion of resistant isolates with the resistance mutation per country, excluding countries with less than 10 resistant isolates for a particular drug. We computed the standard deviation of this distribution of proportions across countries. To test cross-country differences in the proportion of INH phenotypically resistant isolates that contained specific mutations, we used a Fisher-test using the python library scipy¹⁵.

GDP/Programmatic Spending

To correlate countries' gross domestic product (GDP) per capita against resistance acquisition dates, GDP per capita data for 2019 was gathered from the International Monetary Fund (IMF)¹⁶ and plotted against the median MRSCA date for RIF using Microsoft Excel version 16.25. F-value and significance was calculated via Anova¹⁷ in Excel.

Supplementary Material

Supplementary Table 1 Global Lineage Distribution

	Lineage-1	Lineage-2	Lineage-3	Lineage-4	Lineage-5	Lineage-6	Lineage-7	Total
Belarus	0	87	0	46	0	0	0	135
Canada	0	0	0	17	0	0	0	165
China	0	132	2	35	0	0	0	170
Germany	0	22	0	1	0	0	0	857
India	5	9	5	0	0	0	0	52
Iran	0	16	3	1	0	0	0	22
Malawi	208	10	162	945	0	0	0	1427
Mali	0	0	0	37	0	0	0	37
Moldova	0	16	0	35	0	0	0	51
Netherlands	7	19	0	70	0	0	0	98
Peru	0	82	0	869	0	0	0	1098
Portugal	0	2	0	10	0	0	0	13
Romania	0	0	0	33	0	0	0	33
Russia	0	525	1	273	0	0	0	868
Sierra Leone	3	2	0	41	4	10	0	79
South Africa	8	144	23	427	0	0	0	974
South Korea	0	37	1	2	0	0	0	80
Swaziland	0	9	0	0	0	0	0	10
Sweden	2	11	4	11	0	0	0	28
Thailand	0	2	0	0	0	0	0	17
Turkmenistan	0	10	0	0	0	0	0	11
Uganda	2	3	18	50	0	0	0	80
United Kingdom	222	102	560	716	9	4	1	1873
United States of America	1	3	0	0	0	0	0	34
Uzbekistan	0	20	0	0	0	0	0	265
Unknown	37	81	35	792	1	0	0	1749
Total	495	1344	814	4411	14	14	1	10226

Supplementary Table 2 Global Resistance Distribution and Other Resistance Count

Abbreviations: Susceptible (S), Multi-drug Resistant (MDR), Extensively Drug Resistant (XDR), Isoniazid Mono Resistant (INH_MONO), Streptomycin Mono Resistant (STR_MONO), Other Resistant (OTHER_R), Sum With Data (SUM_W_DATA)

country	S	MDR	XDR	INH_MONO	STR_MONO	OTHER_R	NO_DATA	SUM	SUM_W_DATA	MDR/OTHER_R
Belarus	28	1	0	1	0	0	105	135	30	NA
Canada	163	0	0	2	0	0	0	165	165	NA
China	46	117	23	0	0	0	7	170	163	NA
Germany	749	13	0	40	34	14	7	857	850	0.928571429
India	22	13	0	2	2	0	13	52	39	NA
Iran	15	6	0	0	0	0	1	22	21	NA
Malawi	1302	7	0	90	1	3	24	1427	1403	2.333333333
Mali	20	9	0	6	0	0	2	37	35	NA
Moldova	38	2	0	0	0	0	11	51	40	NA
Netherlands	68	0	0	5	20	0	5	98	93	NA
Not Provided	179	646	138	90	6	109	719	1749	1030	5.926605505
Peru	94	674	86	34	5	134	157	1098	941	5.029850746
Romania	0	33	6	0	0	0	0	33	33	NA
Russia	261	420	42	80	17	76	14	868	854	5.526315789
Sierra Leone	41	8	0	7	13	10	0	79	79	0.8
South Africa	612	172	102	34	8	56	92	974	882	3.071428571
South Korea	16	25	21	0	0	0	39	80	41	NA
Swaziland	8	1	0	0	0	1	0	10	10	1
Thailand	1	15	0	0	0	0	1	17	16	NA
Turkmenistan	3	1	0	3	1	3	0	11	11	0.333333333
Uganda	10	45	0	1	0	2	22	80	58	22.5
United Kingdom	1525	83	4	187	5	50	23	1873	1850	1.66
United States of America	0	29	1	0	0	1	4	34	30	29
Uzbekistan	0	263	3	1	0	1	0	265	265	263
Total	5201	2583	426	583	112	460	1246	10185	8939	5.615217391

Abbreviations: RIF (Rifamycin), ETH (Ethionamide), FQ (Fluoroquinolones), SLIS (Second Line Injectable), “_” (AND—EX: RIF_ISONIAZID_ETH = Isolate resistant to Rifamycin, Isoniazid, and Ethionamide)

Other Resistance Categories	Count
RIF_	78
PYRAZINAMIDE_	65
ETHAMBUTOL_	44
ISONIAZID_ETHAMBUTOL_STREPTOMYCIN_	29
ISONIAZID_ETHAMBUTOL_	28
ISONIAZID_PYRAZINAMIDE_	21
ISONIAZID_PYRAZINAMIDE_STREPTOMYCIN_	16
ISONIAZID_ETHAMBUTOL_PYRAZINAMIDE_STREPTOMYCIN_	11
RIF_ETHAMBUTOL_	11
RIF_STREPTOMYCIN_	10
ISONIAZID_STREPTOMYCIN_ETH_	9
ISONIAZID_ETHAMBUTOL_FQ_	9
SLIS_	8
ISONIAZID_STREPTOMYCIN_FQ_	8
ETHAMBUTOL_FQ_	7
FQ_	7
ISONIAZID_ETHAMBUTOL_STREPTOMYCIN_FQ_	7
ETHAMBUTOL_STREPTOMYCIN_	6
ISONIAZID_ETH_	6
STREPTOMYCIN_FQ_	5
ISONIAZID_ETHAMBUTOL_PYRAZINAMIDE_ETH_	5
ISONIAZID_FQ_	4
ISONIAZID_ETHAMBUTOL_STREPTOMYCIN_ETH_	4
ISONIAZID_ETHAMBUTOL_PYRAZINAMIDE_STREPTOMYCIN_SLIS_ETH_	3
ISONIAZID_ETHAMBUTOL_PYRAZINAMIDE_STREPTOMYCIN_SLIS_	3
ISONIAZID_PYRAZINAMIDE_ETH_	3
ETHAMBUTOL_PYRAZINAMIDE_	3
RIF_PYRAZINAMIDE_STREPTOMYCIN_	2
RIF_ETHAMBUTOL_PYRAZINAMIDE_STREPTOMYCIN_FQ_	2
ISONIAZID_ETHAMBUTOL_PYRAZINAMIDE_STREPTOMYCIN_ETH_	2
PYRAZINAMIDE_STREPTOMYCIN_	2
RIF_ETHAMBUTOL_STREPTOMYCIN_	2
ISONIAZID_SLIS_	2
PYRAZINAMIDE_SLIS_	2
ISONIAZID_ETHAMBUTOL_PYRAZINAMIDE_SLIS_ETH_	2
ISONIAZID_ETHAMBUTOL_PYRAZINAMIDE_	2
PYRAZINAMIDE_FQ_	1
ISONIAZID_PYRAZINAMIDE_FQ_SLIS_	1
STREPTOMYCIN_SLIS_	1
RIF_ETHAMBUTOL_FQ_	1
ISONIAZID_ETHAMBUTOL_STREPTOMYCIN_SLIS_ETH_	1
ISONIAZID_ETHAMBUTOL_STREPTOMYCIN_SLIS_	1
RIF_ETHAMBUTOL_STREPTOMYCIN_SLIS_ETH_	1
ETH_	1
RIF_PYRAZINAMIDE_FQ_	1
RIF_ETHAMBUTOL_PYRAZINAMIDE_FQ_	1
RIF_ETHAMBUTOL_PYRAZINAMIDE_STREPTOMYCIN_ETH_	1
RIF_PYRAZINAMIDE_	1
ETHAMBUTOL_ETH_	1
ISONIAZID_ETHAMBUTOL_ETH_	1
RIF_ETHAMBUTOL_SLIS_	1
ISONIAZID_PYRAZINAMIDE_STREPTOMYCIN_ETH_	1
RIF_PYRAZINAMIDE_SLIS_ETH_	1
RIF_ETHAMBUTOL_PYRAZINAMIDE_	1
RIF_ETHAMBUTOL_PYRAZINAMIDE_STREPTOMYCIN_	1
ISONIAZID_ETHAMBUTOL_FQ_SLIS_	1
ETHAMBUTOL_STREPTOMYCIN_FQ_	1
PYRAZINAMIDE_SLIS_ETH_PAS_	1
ISONIAZID_STREPTOMYCIN_FQ_SLIS_ETH_PAS_	1
ISONIAZID_STREPTOMYCIN_SLIS_PAS_	1
ISONIAZID_ETHAMBUTOL_SLIS_	1
RIF_ETHAMBUTOL_PYRAZINAMIDE_FQ_SLIS_ETH_	1
RIF_FQ_SLIS_	1
ISONIAZID_PYRAZINAMIDE_SLIS_	1
ISONIAZID_ETHAMBUTOL_PYRAZINAMIDE_SLIS_ETH_PAS_	1
RIF_ETHAMBUTOL_PYRAZINAMIDE_SLIS_	1
RIF_PYRAZINAMIDE_STREPTOMYCIN_FQ_SLIS_	1
ISONIAZID_ETHAMBUTOL_PYRAZINAMIDE_SLIS_	1
RIF_SLIS_	1
RIF_FQ_	1

Supplementary Table 3 Pairwise Country MRSCA date comparison

Dark red indicates P-value <0.001 (Bonferroni threshold); pink indicates P-value <0.01; white indicates P-value ≥0.01. *No PZA resistance phenotypes was available on Chinese isolates.

INH	Peru	South Africa	China	UK	Russia
Peru		6 E-9	0.1	0.2	9 E-5
South Africa			7 E-29	2 E-32	3 E-7
China				0.2	4 E-17
UK					5 E-15
Russia					

FLQ	Peru	South Africa	China	UK	Russia
Peru		0.3	6 E-10	1 E-6	1 E-4
South Africa			2 E-19	1 E-15	2 E-7
China				0.04	0.6
UK					0.7
Russia					

SLIs	Peru	South Africa	China	UK	Russia
Peru		0.4	1 E-8	2 E-5	0.2
South Africa			5 E-12	9 E-8	0.5
China				0.8	2 E-4
UK					1 E-3
Russia					

RIF	Peru	South Africa	China	UK	Russia
Peru		3 E-9	1 E-14	2 E-8	0.9
South Africa			1 E-31	4 E-21	1 E-3
China				0.02	3 E-17
UK					3 E-11
Russia					

PZA*	Peru	South Africa	China	UK	Russia
Peru		0.01		8 E-9	1 E-4
South Africa				2 E-11	4 E-3
China					
UK					3 E-4
Russia					

EMB	Peru	South Africa	China	UK	Russia
Peru		2 E-9	1 E-7	7 E-6	0.9
South Africa			2 E-17	1 E-12	1 E-6
China				0.4	3 E-7
UK					3 E-6
Russia					

Supplementary Table 4 MDR Frequency Comparison:

Comparison of MDR Frequency between our estimate based on WGS data (WGS MDR Frequency) and estimate based on WHO data (WHO MDR Frequency). Number of isolates with resistance data per country is provided in the “Total Number of Isolates” column.

	WGS MDR Frequency	WHO MDR Frequency	Total Number of Isolates
Belarus	0.03±0.07	0.71(0.67-0.75)	30
Canada	0	0.01(0.01-0.02)	165
China	0.72±0.07	0.08(0.07-0.09)	163
Germany	0.02±0.01	0.03(0.02-0.05)	850
India	0.33±0.15	0.05(0.04-0.06)	39
Iran	0.29±0.19	0.02(0.01-0.02)	21
Malawi	0.01±0.01	0.01(0.01-0.02)	1403
Mali	0.26±0.14	0.03(0.02-0.04)	35
Moldova	0.05±0.07	0.34(0.31-0.36)	40
Netherlands	0	0.02(<0.01-0.01)	93
Not Provided	0.63±0.03		1030
Peru	0.70±0.03	0.09 (0.09-0.1)	941
Romania	1	0.05(0.05-0.06)	33
Russia	0.49±0.03	0.43(0.43-0.44)	854
Sierra Leone	0.10±0.07	0.03(0.02-0.04)	79
South Africa	0.20±0.03	0.04(0.04-0.05)	882
South Korea	0.61±0.15	0.04(0.03-0.04)	41
Swaziland	0.10±0.19	0.1(0.09-0.11)	10
Thailand	0.94±0.12	0.03(0.03-0.04)	16
Turkmenistan	0.10±0.17	0.22(0.19-0.24)	11
Uganda	0.78±0.11	0.02(0.02-0.03)	58
United Kingdom	0.04±0.01	0.02(0.01-0.02)	1850
United States of America	0.97±0.06	0.02(0.01-0.02)	30
Uzbekistan	0.99±0.01	0.33(0.29-0.34)	265

Supplementary Table 5 Dates Drugs Introduced

Drug	Year Introduced
Isoniazid ¹⁸	1952
Rifampicin ¹⁸	1965
Pyrazinamide ¹⁹	1952
Streptomycin ¹⁸	1944
Ethambutol ¹⁸	1965
Ethionamide ¹⁸	1960
Kanamycin/amikacin ¹⁸	1958
Cycloserine ¹⁸	1955
Capreomycin ¹⁸	1967
PAS ¹⁸	1944
Ofloxacin ²⁰	1985
Moxifloxacin ²⁰	1999
Ciprofloxacin ²⁰	1986

Supplementary Table 6 Resistance Mutation Counts Per Country and Per Isolate Phenotype

Amikacin (AMK) Data, Abbreviations: counts = how many isolates had this mutation, no_data = how many isolates had this mutation but no data on whether they were phenotypically resistant or not to AMK, susceptible = how many isolates had this mutation but were phenotypically susceptible to AMK, resistant = how many isolates had this mutation but were resistant to AMK

country	AMK_SNP_N_1472359_A514C_rrs	AMK_SNP_N_1473246_A1401G_rrs	AMK_othersnp	AMK_rrs_A906G_u	AMK_rrs_C517T_u	AMK_rrs_G1484T_u	AMK_unknown
Azerbaijan	0	0	0	0	0	0	0
Bangladesh	0	0	0	0	0	0	0
Belarus	20	51	0	0	0	0	0
Brazil	0	0	0	0	0	0	0
Burma	0	0	0	0	0	0	0
Canada	0	0	0	0	0	0	0
China	10	14	0	0	0	0	0
Colombia	0	1	0	0	0	0	0
Democratic Republic of the Congo	0	0	0	0	0	0	0
Denmark	0	0	0	0	0	0	0
Djibouti	0	0	0	0	0	0	0
Dominican Republic	0	0	0	0	0	0	0
Estonia	0	0	0	0	0	0	0
Georgia	0	3	0	0	0	0	0
Germany	3	3	0	0	0	0	1
Guinea	0	0	0	0	0	0	0
India	0	0	0	0	0	0	0
Indonesia	0	0	0	0	0	0	0
Iran	1	3	0	0	0	0	0
Kazakhstan	0	0	0	0	0	0	0
Malawi	0	0	0	0	0	0	0
Mali	1	0	0	0	0	0	0
Moldova	0	0	0	0	0	0	0
Morocco	0	0	0	0	0	0	0
Nepal	1	0	0	0	0	0	0
Netherlands	1	0	0	0	0	0	0
Nigeria	0	0	0	0	0	0	0
Not Provided	39	160	5	0	8	4	100
Pakistan	0	0	0	0	0	0	0
Peru	14	110	6	3	0	0	48
Philippines	0	0	0	0	0	0	0
Portugal	0	0	0	0	0	0	0
Romania	2	12	0	0	0	0	0
Russia	17	35	0	0	14	1	15
Rwanda	0	2	0	0	0	0	0
Sierra Leone	0	0	0	0	0	0	0
South Africa	83	164	0	0	0	0	11
South Korea	2	26	0	1	0	0	23
Spain	0	0	0	0	0	0	0
Swaziland	0	0	0	0	0	0	0
Thailand	0	0	0	0	0	0	0
Turkmenistan	0	0	0	0	0	0	0
Uganda	0	1	0	0	0	0	0
United Kingdom	16	8	0	0	0	0	0
Uzbekistan	13	46	1	0	0	0	15
Vietnam	0	0	0	0	0	0	0

resistance mutation	counts	no_data	susceptible	resistant
AMK_SNP_N_1473246_A1401G_rrs	639	208	45	386
AMK_SNP_N_1472359_A514C_rrs	223	84	38	101
AMK_unknown	213	0	0	213
AMK_rrs_C517T_u	22	0	0	22
AMK_rrs_G1484T_u	5	0	0	5
AMK_rrs_A906G_u	4	0	0	4
AMK_rrs_C513T_u	4	0	0	4
AMK_rrs_A1461G_u	1	0	0	1
AMK_rrs_A908C_u	1	0	0	1
AMK_rrs_C1105G_u	1	0	0	1
AMK_rrs_C1402T_u	1	0	0	1
AMK_rrs_C799A_u	1	0	0	1
AMK_rrs_C905A_u	1	0	0	1
AMK_rrs_C905G_u	1	0	0	1
AMK_rrs_G878A_u	1	0	0	1

Capreomycin (CAP) Data, Abbreviations: counts = how many isolates had this mutation, no_data = how many isolates had this mutation but no data on whether they were phenotypically resistant or not to CAP, susceptible = how many isolates had this mutation but were phenotypically susceptible to CAP, resistant = how many isolates had this mutation but were resistant to CAP

country	CAP SNP N 1472753 A908C rrs	CAP SNP N 1473246 A1401G rrs	CAP othersnp	CAP rrs A514C u	CAP rrs C517T u	CAP rrs G1484T u	CAP unknown
Azerbaijan	0	0	0	0	0	0	0
Bangladesh	0	0	0	0	0	0	0
Belarus	0	51	0	0	0	0	0
Brazil	0	0	0	0	0	0	0
Burma	0	0	0	0	0	0	0
Canada	0	0	0	0	0	0	0
China	0	14	0	1	1	0	9
Colombia	0	1	0	0	0	0	0
Democratic Republic of the Congo	0	0	0	0	0	0	0
Denmark	0	0	0	0	0	0	0
Djibouti	0	0	0	0	0	0	0
Dominican Republic	0	0	0	0	0	0	0
Estonia	0	0	0	0	0	0	0
Georgia	0	3	0	0	0	0	0
Germany	0	3	0	0	0	0	1
Guinea	0	0	0	0	0	0	0
India	0	0	0	0	0	0	0
Indonesia	0	0	0	0	0	0	0
Iran	0	3	0	0	0	0	0
Kazakhstan	0	0	0	0	0	0	0
Malawi	0	0	0	0	0	0	0
Mali	0	0	0	0	0	0	0
Moldova	0	0	0	0	0	0	0
Morocco	0	0	0	0	0	0	0
Nepal	0	0	0	0	0	0	0
Netherlands	0	0	0	0	0	0	0
Nigeria	0	0	0	0	0	0	0
Not Provided	2	160	7	1	0	7	53
Pakistan	0	0	0	0	0	0	0
Peru	3	110	13	4	0	0	87
Philippines	0	0	0	0	0	0	0
Portugal	0	0	0	0	0	0	0
Romania	1	12	0	0	0	0	0
Russia	1	35	0	0	21	1	20
Rwanda	0	2	0	0	0	0	0
Sierra Leone	0	0	0	0	0	0	0
South Africa	0	164	0	1	0	0	15
South Korea	0	26	1	0	0	0	9
Spain	0	0	0	0	0	0	0
Swaziland	0	0	0	0	0	0	0
Thailand	0	0	0	0	0	0	0
Turkmenistan	0	0	0	0	0	0	0
Uganda	4	1	0	0	0	0	0
United Kingdom	0	8	0	0	0	0	1
Uzbekistan	0	46	2	1	0	0	13
Vietnam	0	0	0	0	0	0	0

resistance mutation	counts	no_data	susceptible	resistant
CAP_SNP_N_1473246_A1401G_rrs	639	167	127	345
CAP_unknown	208	0	0	208
CAP_rrs_C517T_u	22	0	0	22
CAP_SNP_N_1472753_A908C_rrs	11	4	7	0
CAP_rrs_A514C_u	8	0	0	8
CAP_rrs_G1484T_u	8	0	0	8
CAP_rrs_A906G_u	5	0	0	5
CAP_rrs_C513T_u	5	0	0	5
CAP_rrs_C1402T_u	4	0	0	4
CAP_rrs_C1105G_u	1	0	0	1
CAP_rrs_C1390T_u	1	0	0	1
CAP_rrs_C1402A_u	1	0	0	1
CAP_rrs_C239A_u	1	0	0	1
CAP_rrs_C905A_u	1	0	0	1
CAP_rrs_T1322G_u	1	0	0	1
CAP_rrs_T16A_u	1	0	0	1
CAP_rrs_T16C_u	1	0	0	1
CAP_rrs_T556C_u	1	0	0	1
CAP_SNP_N_1473109_T1264G_rrs	0	0	0	0
CAP_SNP_N_1473160_G1315A_rrs	0	0	0	0
CAP_SNP_N_1473343_G1498T_rrs	0	0	0	0

Ciprofloxacin (CIP) Data, Abbreviations: counts = how many isolates had this mutation, no_data = how many isolates had this mutation but no data on whether they were phenotypically resistant or not to CIP, susceptible = how many isolates had this mutation but were phenotypically susceptible to CIP, resistant = how many isolates had this mutation but were resistant to CIP

country	CIP SNP CN 7570 CT gyrA A90V	CIP SNP CN 7572 TC gyrA S91P	CIP SNP CN 7581 GT gyrA D94Y	CIP SNP CN 7582 AC gyrA D94A	CIP SNP CN 7582 AG gyrA D94G	CIP othersnp	CIP unknown
Azerbaijan	0	0	0	0	0	0	0
Bangladesh	0	0	0	0	0	0	0
Belarus	23	6	3	11	21	2	0
Brazil	0	0	0	0	0	0	0
Burma	0	0	0	1	0	0	0
Canada	0	0	0	0	0	0	0
China	11	4	4	7	18	0	0
Colombia	0	0	0	0	0	0	0
Democratic Republic of the Congo	0	0	0	0	0	0	0
Denmark	0	0	0	0	0	0	0
Djibouti	0	0	0	0	0	0	0
Dominican Republic	0	0	0	0	0	0	0
Estonia	0	0	0	0	1	0	0
Georgia	0	0	0	0	0	0	0
Germany	0	2	0	1	2	0	0
Guinea	0	0	0	0	0	0	0
India	2	0	1	0	2	0	0
Indonesia	0	0	0	0	0	0	0
Iran	3	0	0	1	6	0	0
Kazakhstan	0	0	0	0	0	0	0
Malawi	0	0	0	0	0	0	0
Mali	1	0	0	0	0	0	0
Moldova	0	0	0	0	0	1	0
Morocco	0	0	0	0	0	0	0
Nepal	0	0	0	0	0	0	0
Netherlands	0	0	0	0	1	0	0
Nigeria	0	0	0	0	0	0	0
Not Provided	57	36	13	25	86	11	6
Pakistan	1	0	0	0	0	0	0
Peru	20	1	2	9	29	7	0
Philippines	0	0	0	0	0	0	0
Portugal	0	0	0	0	0	0	0
Romania	2	0	1	3	1	0	0
Russia	10	5	3	12	31	0	0
Rwanda	1	0	0	0	0	0	0
Sierra Leone	0	0	0	0	0	0	0
South Africa	78	4	6	21	44	1	0
South Korea	13	3	1	2	14	4	22
Spain	0	0	0	0	0	0	0
Swaziland	0	0	0	0	0	0	0
Thailand	0	0	0	0	0	0	0
Turkmenistan	0	0	0	0	0	0	0
Uganda	0	1	0	0	0	0	0
United Kingdom	7	7	0	1	30	2	0
Uzbekistan	0	0	0	0	0	0	0
Vietnam	0	0	0	0	0	0	0

resistance mutation	counts	no_data	susceptible	resistant
CIP_SNP_CN_7582_AG_gyrA_D94G	286	249	1	36
CIP_SNP_CN_7570_CT_gyrA_A90V	229	201	12	16
CIP_SNP_CN_7582_AC_gyrA_D94A	94	92	2	0
CIP_SNP_CN_7572_TC_gyrA_S91P	69	63	2	4
CIP_SNP_CN_7581_GT_gyrA_D94Y	34	28	0	6
CIP_unknown	28	0	0	28
CIP_SNP_CN_6735_AC_gyrB_N538T	10	10	0	0
CIP_SNP_CN_7566_GA_gyrA_D89N	6	6	0	0
CIP_gyrA_A74S_u	2	0	0	2
CIP_gyrA_D94N_u	2	0	0	2
CIP_gyrB_A471V_u	2	0	0	2
CIP_gyrB_S486F_u	2	0	0	2
CIP_gyrB_T539N_u	2	0	0	2
CIP_gyrA_A288D_u	1	0	0	1
CIP_gyrA_D94H_u	1	0	0	1

Ethambutol (EMB) Data, Abbreviations: counts = how many isolates had this mutation, no_data = how many isolates had this mutation but no data on whether they were phenotypically resistant or not to EMB, susceptible = how many isolates had this mutation but were phenotypically susceptible to EMB, resistant = how many isolates had this mutation but were resistant to EMB

country	EMB_SNP_CN_4242182_GT_embC_A774S	EMB_SNP_CN_4247429_AG_embB_M306V	EMB_SNP_CN_4247431_GA_embB_M306I	EMB_SNP_CN_4247431_GC_embB_M306I	EMB_SNP_CN_4247431_GT_embB_M306I	EMB_othersnp	EMB_unknown
Azerbaijan	0	0	0	0	0	1	0
Bangladesh	0	0	0	0	0	1	2
Belarus	0	16	46	46	46	30	0
Brazil	0	0	0	0	0	0	0
Burma	0	0	1	1	1	1	0
Canada	0	0	0	0	0	0	0
China	0	28	17	17	17	45	5
Colombia	0	0	0	0	0	0	0
Democratic Republic of the Congo	0	0	0	0	0	0	0
Denmark	0	0	0	0	0	0	0
Djibouti	0	0	0	0	0	0	0
Dominican Republic	0	0	0	0	0	0	0
Estonia	0	2	0	0	0	0	0
Georgia	0	0	0	0	0	4	0
Germany	3	4	4	4	4	12	1
Guinea	0	1	0	0	0	0	0
India	0	4	0	0	0	4	0
Indonesia	0	0	0	0	0	0	0
Iran	0	4	2	2	2	6	1
Kazakhstan	0	2	0	0	0	0	0
Malawi	2	1	0	0	0	5	0
Mali	0	8	4	4	4	3	0
Moldova	0	5	5	5	5	11	0
Morocco	0	0	0	0	0	1	0
Nepal	0	0	0	0	0	0	1
Netherlands	0	0	0	0	0	0	0
Nigeria	0	0	1	1	1	0	0
Not Provided	54	203	135	135	135	280	34
Pakistan	0	1	0	0	0	0	0
Peru	220	139	178	178	178	367	121
Philippines	0	0	0	0	0	0	1
Portugal	0	0	1	1	1	0	0
Romania	0	6	11	11	11	17	0
Russia	0	93	46	46	46	232	46
Rwanda	0	1	0	0	0	1	1
Sierra Leone	1	1	4	4	4	7	2
South Africa	0	139	132	132	132	56	6
South Korea	0	11	12	12	12	16	4
Spain	0	1	0	0	0	0	0
Swaziland	0	0	0	0	0	1	0
Thailand	0	0	0	0	0	0	0
Turkmenistan	0	1	1	1	1	0	0
Uganda	1	7	18	18	18	18	1
United Kingdom	10	21	45	45	45	31	1
Uzbekistan	0	115	56	56	56	60	7
Vietnam	0	0	0	0	0	0	0

resistance mutation	counts	no_data	susceptible	resistant
EMB SNP CN 4247429 AG embB M306V	814	123	185	506
EMB SNP CN 4247431 GA embB M306I	719	122	185	412
EMB SNP CN 4247431 GC embB M306I	719	122	185	412
EMB SNP CN 4247431 GT embB M306I	719	122	185	412
EMB SNP CN 4242182 GT embC A774S	291	33	89	169
EMB SNP CN 4247574 AC embB D354A	236	6	108	122
EMB unknown	234	0	0	234
EMB SNP CN 4248003 AG embB Q497R	164	42	40	82
EMB SNP CN 4247730 GC embB G406A	129	20	33	76
EMB SNP CN 409569 GA iniB A70T	117	19	25	73
EMB SNP CN 4247729 GA embB G406S	71	9	14	48
EMB SNP CN 4249583 GA embB D1024N	56	7	17	32
EMB SNP CN 4243392 AG embA N54D	50	8	1	41
EMB SNP P 4243222 CA.11 embA.embB	41	11	7	23
EMB SNP CN 4247429 AC embB M306L	34	7	8	19
EMB embB G406D u	31	0	0	31
EMB SNP P 4243225 CT.8 embA.embB	28	4	3	21
EMB upstream intergenic-embA-embB C16G u	22	0	0	22
EMB SNP CN 4249518 AG embB H1002R	19	0	4	15
EMB SNP CN 4247495 GT embB D328Y	17	3	2	12
EMB upstream intergenic-embA-embB C12T u	17	0	0	17
EMB embB D328G u	12	0	0	12
EMB embB L74R u	12	0	0	12
EMB embB Q497K u	12	0	0	12
EMB upstream intergenic-embA-embB C16T u	10	0	0	10
EMB embB E405D u	8	0	0	8
EMB embB N296H u	8	0	0	8
EMB embB S297A u	7	0	0	7
EMB upstream intergenic-embA-embB G43C u	7	0	0	7
EMB SNP CN 4247729 GT embB G406C	6	3	1	2
EMB embB A409P u	5	0	0	5
EMB upstream intergenic-embA-embB C16A u	5	0	0	5
EMB embB Q445R u	4	0	0	4
EMB embB Q497P u	4	0	0	4
EMB embB Y319C u	4	0	0	4
EMB embB Y319S u	4	0	0	4
EMB embB G603R u	3	0	0	3
EMB embB I450L u	3	0	0	3
EMB embB T1082A u	3	0	0	3
EMB embB T642A u	3	0	0	3
EMB embA D808A u	2	0	0	2
EMB embA I905V u	2	0	0	2
EMB embA L535V u	2	0	0	2
EMB embA S654L u	2	0	0	2
EMB embA V534G u	2	0	0	2
EMB embB E504D u	2	0	0	2
EMB embB S347I u	2	0	0	2
EMB embB V50A u	2	0	0	2
EMB embC M351T u	2	0	0	2
EMB iniB A182V u	2	0	0	2
EMB upstream intergenic-embA-embB G76C u	2	0	0	2
EMB embA G554D u	1	0	0	1
EMB embA G884D u	1	0	0	1
EMB embA T652R u	1	0	0	1
EMB embA V31I u	1	0	0	1
EMB embA V391L u	1	0	0	1
EMB embA V479M u	1	0	0	1
EMB embB C361Y u	1	0	0	1
EMB embB D1056A u	1	0	0	1
EMB embB D300G u	1	0	0	1
EMB embB F323L u	1	0	0	1
EMB embB I1009L u	1	0	0	1
EMB embB I72L u	1	0	0	1
EMB embB I72S u	1	0	0	1
EMB embB L402V u	1	0	0	1
EMB embB M306T u	1	0	0	1
EMB embB M396I u	1	0	0	1
EMB embB N399T u	1	0	0	1
EMB embB N644I u	1	0	0	1
EMB embB P404S u	1	0	0	1
EMB embB P731L u	1	0	0	1
EMB embB S347T u	1	0	0	1
EMB embB S412P u	1	0	0	1
EMB embB S538P u	1	0	0	1
EMB embB T546I u	1	0	0	1
EMB embB T581A u	1	0	0	1
EMB embB T643I u	1	0	0	1
EMB embB W332R u	1	0	0	1
EMB embB Y319D u	1	0	0	1
EMB embC A68T u	1	0	0	1
EMB embC C411Y u	1	0	0	1
EMB embC I10T u	1	0	0	1
EMB iniB G386D u	1	0	0	1
EMB upstream intergenic-embA-embB C15G u	1	0	0	1
EMB upstream intergenic-embA-embB d42CAT u	1	0	0	1
EMB upstream intergenic-embA-embB d43G u	1	0	0	1
EMB upstream intergenic-embA-embB 1.14-5TACCATCGAG u	1	0	0	1

Ethionamide (ETH) Data, Abbreviations: counts = how many isolates had this mutation, no_data = how many isolates had this mutation but no data on whether they were phenotypically resistant or not to ETH, susceptible = how many isolates had this mutation but were phenotypically susceptible to ETH, resistant = how many isolates had this mutation but were resistant to ETH

country	ETH SNP_CN_1674263_TC_inhA_I21T	ETH SNP_CN_1674481_TG_inhA_S94A	ETH SNP_CN_4326333_CG_ethA_A381P	ETH SNP_P_1673423_GT.17_fabG1.inhA	ETH SNP_P_1673425_CT.15_fabG1.inhA	ETH_othersnp	ETH_unknown
Azerbaijan	0	0	0	0	0	0	0
Bangladesh	0	0	0	0	0	0	0
Belarus	0	0	0	0	30	0	0
Brazil	0	0	0	0	0	0	0
Burma	0	0	0	0	0	0	0
Canada	0	0	0	0	0	0	0
China	0	1	0	0	18	13	15
Colombia	0	0	0	1	0	0	0
Democratic Republic of the Congo	0	0	0	0	0	0	0
Denmark	0	0	0	0	0	0	0
Djibouti	0	0	0	0	0	0	0
Dominican Republic	0	0	0	0	0	0	0
Estonia	0	0	0	0	0	0	0
Georgia	0	0	0	0	1	0	0
Germany	0	3	0	0	21	0	0
Guinea	0	0	0	0	2	0	0
India	0	0	0	0	1	0	0
Indonesia	0	0	0	0	0	0	0
Iran	0	0	0	0	0	0	0
Kazakhstan	0	0	0	0	0	0	0
Malawi	1	11	0	0	15	0	0
Mali	0	1	0	0	2	0	0
Moldova	0	0	0	0	25	0	0
Morocco	0	0	0	0	0	0	0
Nepal	0	0	0	0	0	0	0
Netherlands	0	0	0	0	0	0	0
Nigeria	0	0	0	0	0	0	0
Not Provided	12	98	12	14	573	51	46
Pakistan	0	0	0	0	1	0	0
Peru	33	20	1	27	178	100	76
Philippines	0	0	0	0	0	0	0
Portugal	0	0	0	0	0	0	0
Romania	0	0	0	0	7	0	0
Russia	0	4	0	2	50	0	3
Rwanda	0	0	0	0	1	0	0
Sierra Leone	0	1	0	1	1	0	0
South Africa	8	2	76	37	113	11	6
South Korea	0	2	0	0	8	6	10
Spain	0	1	0	0	1	0	0
Swaziland	0	0	0	0	0	0	0
Thailand	0	0	0	0	0	0	0
Turkmenistan	0	0	0	0	0	0	0
Uganda	0	0	0	0	4	0	1
United Kingdom	2	4	1	6	76	1	0
Uzbekistan	1	0	0	1	9	0	1
Vietnam	0	0	0	0	1	0	0

resistance mutation	counts	no data	susceptible	resistant
ETH SNP P 1673425 CT.15 fabG1.inhA	1138	716	157	265
ETH unknown	158	0	0	158
ETH SNP CN 1674481 TG inhA S94A	148	52	13	83
ETH SNP CN 4326333 CG ethA A381P	90	21	35	34
ETH SNP P 1673423 GT.17 fabG1.inhA	89	32	38	19
ETH SNP CN 1674263 TC inhA I21T	57	18	10	29
ETH SNP CN 4326116 GA ethA T453I	35	6	14	15
ETH SNP CN 4327380 AC ethA Y32D	15	2	9	4
ETH upstream intergenic-fabG1-inhA T8A u	14	0	0	14
ETH DEL F 4326184 d1289G ethA	12	5	4	3
ETH SNP CN 4327416 CA ethA A20S	12	5	4	3
ETH ethA P378L u	9	0	0	9
ETH INS F 4326722 i751C ethA	6	2	1	3
ETH SNP CN 4326305 GA ethA S390F	6	2	1	3
ETH SNP CN 4326713 TG ethA Q254P	6	1	0	5
ETH upstream intergenic-fabG1-inhA T8C u	6	0	0	6
ETH upstream intergenic-fabG1-inhA T8G u	4	0	0	4
ETH ethA S266R u	3	0	0	3
ETH ethA C403R u	2	0	0	2
ETH ethA E400D u	2	0	0	2
ETH ethA F302L u	2	0	0	2
ETH ethA S110W u	2	0	0	2
ETH ethA Y50C u	2	0	0	2
ETH SNP CN 4327311 AG ethA S55P	1	0	1	0
ETH ethA A76V u	1	0	0	1
ETH ethA F320S u	1	0	0	1
ETH ethA F431V u	1	0	0	1
ETH ethA G139C u	1	0	0	1
ETH ethA G139S u	1	0	0	1
ETH ethA G139V u	1	0	0	1
ETH ethA G182S u	1	0	0	1
ETH ethA G423R u	1	0	0	1
ETH ethA G43C u	1	0	0	1
ETH ethA G450D u	1	0	0	1
ETH ethA H166P u	1	0	0	1
ETH ethA L134P u	1	0	0	1
ETH ethA L194P u	1	0	0	1
ETH ethA L205P u	1	0	0	1
ETH ethA L374R u	1	0	0	1
ETH ethA L397R u	1	0	0	1
ETH ethA N379D u	1	0	0	1
ETH ethA P149S u	1	0	0	1
ETH ethA P257A u	1	0	0	1
ETH ethA P257S u	1	0	0	1
ETH ethA P51H u	1	0	0	1
ETH ethA P51L u	1	0	0	1
ETH ethA P51S u	1	0	0	1
ETH ethA Q206E u	1	0	0	1
ETH ethA Q246R u	1	0	0	1
ETH ethA R38P u	1	0	0	1
ETH ethA R404L u	1	0	0	1
ETH ethA R54S u	1	0	0	1
ETH ethA S18R u	1	0	0	1
ETH ethA S208P u	1	0	0	1
ETH ethA S57F u	1	0	0	1
ETH ethA T189K u	1	0	0	1
ETH ethA T342K u	1	0	0	1
ETH ethA T383P u	1	0	0	1
ETH ethA T392R u	1	0	0	1
ETH ethA T44P u	1	0	0	1
ETH ethA T61M u	1	0	0	1
ETH ethA Y235D u	1	0	0	1
ETH fabG1 L85V u	1	0	0	1
ETH inhA I194T u	1	0	0	1
ETH inhA I21V u	1	0	0	1
ETH inhA I95L u	1	0	0	1
ETH upstream intergenic-fabG1-inhA C34T u	1	0	0	1
ETH INS F 4326141 i1332C ethA	0	0	0	0
ETH SNP CN 1673449 AC fabG1 T4P	0	0	0	0
ETH SNP CN 1674434 TG inhA V78G	0	0	0	0
ETH SNP CN 4326278 GT ethA S399.	0	0	0	0
ETH SNP CN 4326600 GA ethA R292.	0	0	0	0
ETH SNP CN 4326714 GA ethA Q254.	0	0	0	0
ETH SNP CN 4327148 CT ethA W109.	0	0	0	0

Isoniazid (INH) Data, Abbreviations: counts = how many isolates had this mutation, no_data = how many isolates had this mutation but no data on whether they were phenotypically resistant or not to INH, susceptible = how many isolates had this mutation but were phenotypically susceptible to INH, resistant = how many isolates had this mutation but were resistant to INH

country	INH SNP CN 1674481 TG inhA S94A	INH SNP CN 2155168 CG katG S315T	INH SNP CN 2518919 GA kasA G269S	INH SNP CN 409569 GA iniB A70T	INH SNP P 1673425 CT.15 fabG1.inhA	INH othersnp	INH unknown
Azerbaijan	0	1	0	0	0	0	0
Bangladesh	0	1	0	0	0	0	2
Belarus	0	101	23	0	30	43	1
Brazil	0	0	0	0	0	3	0
Burma	0	1	0	0	0	0	0
Canada	0	0	0	0	0	0	0
China	1	65	0	0	18	24	4
Colombia	0	0	0	0	0	1	0
Democratic Republic of the Congo	0	0	0	0	0	0	0
Denmark	0	0	0	0	0	0	0
Djibouti	0	0	0	0	0	0	0
Dominican Republic	0	0	0	0	0	0	0
Estonia	0	2	0	0	0	0	0
Georgia	0	5	0	0	1	0	0
Germany	3	43	24	2	21	4	4
Guinea	0	0	0	0	2	0	0
India	0	8	0	0	1	1	0
Indonesia	0	0	0	0	0	0	0
Iran	0	11	0	0	0	0	1
Kazakhstan	0	2	0	0	0	0	0
Malawi	11	50	24	0	15	4	12
Mali	1	23	0	0	2	5	0
Moldova	0	43	3	0	25	0	0
Morocco	0	1	0	0	0	0	0
Nepal	0	4	0	0	0	0	0
Netherlands	0	3	0	0	0	0	1
Nigeria	0	1	0	0	0	0	0
Not Provided	98	539	120	27	573	102	82
Pakistan	0	0	0	0	1	0	0
Peru	20	593	299	88	178	80	40
Philippines	0	0	0	0	0	0	2
Portugal	0	1	0	0	0	0	0
Romania	0	30	1	0	7	5	0
Russia	4	470	61	0	50	20	49
Rwanda	0	1	0	0	1	1	2
Sierra Leone	1	16	8	0	1	3	2
South Africa	2	299	98	0	113	57	13
South Korea	2	35	0	0	8	3	6
Spain	1	1	0	0	1	0	0
Swaziland	0	0	0	0	0	1	0
Thailand	0	0	0	0	0	0	0
Turkmenistan	0	6	0	0	0	0	0
Uganda	0	35	1	0	4	3	2
United Kingdom	4	173	34	0	76	26	13
Uzbekistan	0	226	11	0	9	24	5
Vietnam	0	0	0	0	1	0	0

resistance mutation	counts	no data	susceptible	resistant
INH SNP CN 2155168 CG katG S315T	2790	210	310	2270
INH SNP P 1673425 CT.15 fabG1.inhA	1138	424	100	614
INH SNP CN 2518919 GA kasA G269S	707	58	248	401
INH unknown	241	0	0	241
INH SNP CN 1674481 TG inhA S94A	148	25	10	113
INH SNP CN 409569 GA iniB A70T	117	19	9	89
INH SNP P 1673432 TC.8 fabG1.inhA	102	33	36	33
INH SNP P 1673423 GT.17 fabG1.inhA	89	8	12	69
INH SNP CN 2155168 CT katG S315N	58	5	5	48
INH SNP CN 2155167 GT katG S315R	56	3	5	48
INH SNP P 1673432 TG.8 fabG1.inhA	10	0	1	9
INH inhA I21T u	3	0	0	3
INH katG G279D u	3	0	0	3
INH katG R104Q u	3	0	0	3
INH katG S315G u	3	0	0	3
INH katG W161R u	3	0	0	3
INH katG D735Y u	2	0	0	2
INH katG L141F u	2	0	0	2
INH katG L141S u	2	0	0	2
INH katG N138H u	2	0	0	2
INH katG W191R u	2	0	0	2
INH katG Y337C u	2	0	0	2
INH katG Y413C u	2	0	0	2
INH ahpC P44R u	1	0	0	1
INH inhA K118E u	1	0	0	1
INH iniB A222T u	1	0	0	1
INH iniB G131R u	1	0	0	1
INH iniB G171D u	1	0	0	1
INH iniB G386D u	1	0	0	1
INH iniB S249R u	1	0	0	1
INH kasA A45S u	1	0	0	1
INH kasA R161S u	1	0	0	1
INH katG A109V u	1	0	0	1
INH katG A139P u	1	0	0	1
INH katG A162E u	1	0	0	1
INH katG A614E u	1	0	0	1
INH katG D142G u	1	0	0	1
INH katG D163N u	1	0	0	1
INH katG D189A u	1	0	0	1
INH katG D259Y u	1	0	0	1
INH katG D282G u	1	0	0	1
INH katG D419H u	1	0	0	1
INH katG D487N u	1	0	0	1
INH katG D656A u	1	0	0	1
INH katG D695A u	1	0	0	1
INH katG D94G u	1	0	0	1
INH katG F408L u	1	0	0	1
INH katG G120S u	1	0	0	1
INH katG G182R u	1	0	0	1
INH katG G273S u	1	0	0	1
INH katG G285S u	1	0	0	1
INH katG G297V u	1	0	0	1
INH katG G299S u	1	0	0	1
INH katG G630R u	1	0	0	1
INH katG G699V u	1	0	0	1
INH katG G99R u	1	0	0	1
INH katG K537E u	1	0	0	1
INH katG L159F u	1	0	0	1
INH katG L159P u	1	0	0	1
INH katG L598F u	1	0	0	1
INH katG L76P u	1	0	0	1
INH katG N138D u	1	0	0	1
INH katG P232R u	1	0	0	1
INH katG P325S u	1	0	0	1
INH katG Q461P u	1	0	0	1
INH katG R571H u	1	0	0	1
INH katG S302R u	1	0	0	1
INH katG S315I u	1	0	0	1
INH katG T180K u	1	0	0	1
INH katG T271I u	1	0	0	1
INH katG T475I u	1	0	0	1
INH katG V320L u	1	0	0	1
INH katG V423I u	1	0	0	1
INH katG V442A u	1	0	0	1
INH katG V626E u	1	0	0	1
INH katG V633A u	1	0	0	1
INH katG W191G u	1	0	0	1
INH katG W300C u	1	0	0	1
INH katG W300R u	1	0	0	1
INH katG W300S u	1	0	0	1
INH katG W328L u	1	0	0	1
INH katG W341G u	1	0	0	1
INH katG W438G u	1	0	0	1
INH katG W90R u	1	0	0	1
INH katG Y413H u	1	0	0	1
INH katG Y608F u	1	0	0	1
INH katG Y98C u	1	0	0	1
INH upstream intergenic-fabG1-inhA G102A u	1	0	0	1
INH upstream intergenic-fabG1-inhA T8A u	1	0	0	1
INH SNP CN 2726338 TG ahpC V49G	0	0	0	0

Kanamycin (KAN) Data, Abbreviations: counts = how many isolates had this mutation, no_data = how many isolates had this mutation but no data on whether they were phenotypically resistant or not to KAN, susceptible = how many isolates had this mutation but were phenotypically susceptible to KAN, resistant = how many isolates had this mutation but were resistant to KAN

country	KAN_SNP_N_1473246_A1401G_rrs	KAN_othersnp	KAN_rrs_A514C_u	KAN_rrs_A906G_u	KAN_rrs_C513T_u	KAN_rrs_G1484T_u	KAN_unknown
Azerbaijan	0	0	0	0	0	0	0
Bangladesh	0	0	0	0	0	0	0
Belarus	51	0	0	0	0	0	0
Brazil	0	0	0	0	0	0	0
Burma	0	0	0	0	0	0	0
Canada	0	0	0	0	0	0	0
China	14	1	2	0	0	0	6
Colombia	1	0	0	0	0	0	0
Democratic Republic of the Congo	0	0	0	0	0	0	0
Denmark	0	0	0	0	0	0	0
Djibouti	0	0	0	0	0	0	0
Dominican Republic	0	0	0	0	0	0	0
Estonia	0	0	0	0	0	0	0
Georgia	3	0	0	0	0	0	0
Germany	3	0	0	0	0	0	0
Guinea	0	0	0	0	0	0	0
India	0	0	0	0	0	0	0
Indonesia	0	0	0	0	0	0	0
Iran	3	0	0	0	0	0	0
Kazakhstan	0	0	0	0	0	0	0
Malawi	0	0	0	0	0	0	0
Mali	0	0	0	0	0	0	0
Moldova	0	0	0	0	0	0	0
Morocco	0	0	0	0	0	0	0
Nepal	0	0	0	0	0	0	0
Netherlands	0	0	0	0	0	0	0
Nigeria	0	0	0	0	0	0	0
Not Provided	160	6	3	0	3	7	53
Pakistan	0	0	0	0	0	0	0
Peru	110	4	1	5	4	0	59
Philippines	0	0	0	0	0	0	0
Portugal	0	0	0	0	0	0	0
Romania	12	0	0	0	0	0	2
Russia	35	0	0	0	0	0	0
Rwanda	2	0	0	0	0	0	0
Sierra Leone	0	0	0	0	0	0	0
South Africa	164	2	2	0	0	0	17
South Korea	26	0	0	1	0	0	24
Spain	0	0	0	0	0	0	0
Swaziland	0	0	0	0	0	0	0
Thailand	0	0	0	0	0	0	0
Turkmenistan	0	0	0	0	0	0	0
Uganda	1	0	0	0	0	0	0
United Kingdom	8	0	0	0	0	0	4
Uzbekistan	46	0	0	0	0	0	0
Vietnam	0	0	0	0	0	0	0

resistance mutation	counts	no_data	susceptible	resistant
KAN SNP N 1473246 A1401G rrs	639	208	83	348
KAN unknown	165	0	0	165
KAN rrs A514C u	8	0	0	8
KAN rrs C513T u	7	0	0	7
KAN rrs G1484T u	7	0	0	7
KAN rrs A906G u	6	0	0	6
KAN rrs C517T u	4	0	0	4
KAN rrs C905A u	2	0	0	2
KAN rrs A365G u	1	0	0	1
KAN rrs A514T u	1	0	0	1
KAN rrs A908C u	1	0	0	1
KAN rrs A908G u	1	0	0	1
KAN rrs C1105G u	1	0	0	1
KAN rrs C1402A u	1	0	0	1
KAN rrs G878A u	1	0	0	1
KAN SNP CN 1918745 AG tlyA .269W	0	0	0	0

Levofloxacin (LEVO) Data, Abbreviations: counts = how many isolates had this mutation, no_data = how many isolates had this mutation but no data on whether they were phenotypically resistant or not to LEVO, susceptible = how many isolates had this mutation but were phenotypically susceptible to LEVO, resistant = how many isolates had this mutation but were resistant to LEVO

country	LEVO_SNP_CN_7570_CT_gyrA_A90V	LEVO_SNP_CN_7572_TC_gyrA_S91P	LEVO_SNP_CN_7581_GA_gyrA_D94N	LEVO_SNP_CN_7582_AC_gyrA_D94A	LEVO_SNP_CN_7582_AG_gyrA_D94G	LEVO_othersnp	LEVO_unknown
Azerbaijan	0	0	0	0	0	0	0
Bangladesh	0	0	0	0	0	0	0
Belarus	23	6	2	11	21	5	0
Brazil	0	0	0	0	0	0	0
Burma	0	0	0	1	0	0	0
Canada	0	0	0	0	0	0	0
China	11	4	1	7	18	4	0
Colombia	0	0	0	0	0	0	0
Democratic Republic of the Congo	0	0	0	0	0	0	0
Denmark	0	0	0	0	0	0	0
Djibouti	0	0	0	0	0	0	0
Dominican Republic	0	0	0	0	0	0	0
Estonia	0	0	0	0	1	0	0
Georgia	0	0	0	0	0	0	0
Germany	0	2	1	1	2	0	0
Guinea	0	0	0	0	0	0	0
India	2	0	0	0	2	1	0
Indonesia	0	0	0	0	0	0	0
Iran	3	0	0	1	6	0	0
Kazakhstan	0	0	0	0	0	0	0
Malawi	0	0	0	0	0	0	0
Mali	1	0	0	0	0	0	0
Moldova	0	0	0	0	0	1	0
Morocco	0	0	0	0	0	0	0
Nepal	0	0	0	0	0	0	0
Netherlands	0	0	0	0	1	0	0
Nigeria	0	0	0	0	0	0	0
Not Provided	57	36	12	25	86	19	10
Pakistan	1	0	0	0	0	0	0
Peru	20	1	8	9	29	10	3
Philippines	0	0	0	0	0	0	0
Portugal	0	0	0	0	0	0	0
Romania	2	0	0	3	1	1	0
Russia	10	5	7	12	31	4	0
Rwanda	1	0	0	0	0	0	0
Sierra Leone	0	0	0	0	0	0	0
South Africa	78	4	6	21	44	12	0
South Korea	13	3	2	2	14	1	0
Spain	0	0	0	0	0	0	0
Swaziland	0	0	0	0	0	0	0
Thailand	0	0	0	0	0	0	0
Turkmenistan	0	0	0	0	0	0	0
Uganda	0	1	0	0	0	0	0
United Kingdom	7	7	2	1	30	1	0
Uzbekistan	0	0	0	0	0	0	0
Vietnam	0	0	0	0	0	0	0

resistance mutation	counts	no_data	susceptible	resistant
LEVO_SNP_CN_7582_AG_gyrA_D94G	286	270	0	16
LEVO_SNP_CN_7570_CT_gyrA_A90V	229	214	2	13
LEVO_SNP_CN_7582_AC_gyrA_D94A	94	91	0	3
LEVO_SNP_CN_7572_TC_gyrA_S91P	69	66	0	3
LEVO_SNP_CN_7581_GA_gyrA_D94N	41	39	0	2
LEVO_SNP_CN_7581_GT_gyrA_D94Y	34	33	0	1
LEVO_unknown	13	0	0	13
LEVO_SNP_CN_6735_AC_gyrB_N538T	10	8	1	1
LEVO_SNP_CN_7563_GT_gyrA_G88C	8	7	0	1
LEVO_SNP_CN_7566_GA_gyrA_D89N	6	5	1	0
LEVO_gyrA_R128K_u	1	0	0	1

Ofloxacin (OFLX) Data, Abbreviations: counts = how many isolates had this mutation, no_data = how many isolates had this mutation but no data on whether they were phenotypically resistant or not to OFLX, susceptible = how many isolates had this mutation but were phenotypically susceptible to OFLX, resistant = how many isolates had this mutation but were resistant to OFLX

country	OFLX_SNP_CN_7570_CT_gyrA_A90V	OFLX_SNP_CN_7581_GA_gyrA_D94N	OFLX_SNP_CN_7582_AC_gyrA_D94A	OFLX_SNP_CN_7582_AG_gyrA_D94G	OFLX_gyrA_S91P_u	OFLX_othersnp	OFLX_unknown
Azerbaijan	0	0	0	0	0	0	0
Bangladesh	0	0	0	0	0	0	0
Belarus	23	2	11	21	0	3	0
Brazil	0	0	0	0	0	0	0
Burma	0	0	1	0	0	0	0
Canada	0	0	0	0	0	0	0
China	11	1	7	18	3	6	13
Colombia	0	0	0	0	0	0	0
Democratic Republic of the Congo	0	0	0	0	0	0	0
Denmark	0	0	0	0	0	0	0
Djibouti	0	0	0	0	0	0	0
Dominican Republic	0	0	0	0	0	0	0
Estonia	0	0	0	1	0	0	0
Georgia	0	0	0	0	0	0	0
Germany	0	1	1	2	0	1	0
Guinea	0	0	0	0	0	0	0
India	2	0	0	2	0	0	0
Indonesia	0	0	0	0	0	0	0
Iran	3	0	1	6	0	0	0
Kazakhstan	0	0	0	0	0	0	0
Malawi	0	0	0	0	0	0	0
Mali	1	0	0	0	0	0	0
Moldova	0	0	0	0	0	0	0
Morocco	0	0	0	0	0	0	0
Nepal	0	0	0	0	0	0	0
Netherlands	0	0	0	1	0	0	0
Nigeria	0	0	0	0	0	0	0
Not Provided	57	12	25	86	26	49	97
Pakistan	1	0	0	0	0	0	0
Peru	20	8	9	29	0	3	3
Philippines	0	0	0	0	0	0	0
Portugal	0	0	0	0	0	0	0
Romania	2	0	3	1	0	4	2
Russia	10	7	12	31	5	11	57
Rwanda	1	0	0	0	0	0	0
Sierra Leone	0	0	0	0	0	0	0
South Africa	78	6	21	44	1	22	16
South Korea	13	2	2	14	0	0	0
Spain	0	0	0	0	0	0	0
Swaziland	0	0	0	0	0	0	0
Thailand	0	0	0	0	0	0	0
Turkmenistan	0	0	0	0	0	0	0
Uganda	0	0	0	0	1	1	0
United Kingdom	7	2	1	30	3	1	0
Uzbekistan	0	0	0	0	0	0	7
Vietnam	0	0	0	0	0	0	0

resistance mutation	counts	no data	susceptible	resistant
OFLX SNP CN 7582 AG gyrA D94G	286	89	30	167
OFLX SNP CN 7570 CT gyrA A90V	229	62	75	92
OFLX unknown	195	0	0	195
OFLX SNP CN 7582 AC gyrA D94A	94	19	23	52
OFLX SNP CN 7581 GA gyrA D94N	41	14	5	22
OFLX gyrA S91P u	39	0	0	39
OFLX gyrA D94Y u	22	0	0	22
OFLX SNP CN 7581 GC gyrA D94H	14	3	5	6
OFLX SNP CN 6735 AC gyrB N538T	10	6	1	3
OFLX gyrA G88C u	4	0	0	4
OFLX gyrB D500N u	4	0	0	4
OFLX gyrB V340L u	4	0	0	4
OFLX gyrA Q613E u	3	0	0	3
OFLX gyrB D500H u	3	0	0	3
OFLX gyrA A74S u	2	0	0	2
OFLX gyrA N282K u	2	0	0	2
OFLX gyrA T267I u	2	0	0	2
OFLX gyrB E540D u	2	0	0	2
OFLX gyrB P133L u	2	0	0	2
OFLX gyrB S486F u	2	0	0	2
OFLX gyrA A288D u	1	0	0	1
OFLX gyrA A463S u	1	0	0	1
OFLX gyrA A90G u	1	0	0	1
OFLX gyrA D829E u	1	0	0	1
OFLX gyrA D89N u	1	0	0	1
OFLX gyrA D94V u	1	0	0	1
OFLX gyrA E214D u	1	0	0	1
OFLX gyrA G88A u	1	0	0	1
OFLX gyrA P472S u	1	0	0	1
OFLX gyrA R292G u	1	0	0	1
OFLX gyrA R382L u	1	0	0	1
OFLX gyrA R448H u	1	0	0	1
OFLX gyrA R592S u	1	0	0	1
OFLX gyrA T135S u	1	0	0	1
OFLX gyrB A471V u	1	0	0	1
OFLX gyrB A543T u	1	0	0	1
OFLX gyrB H350Y u	1	0	0	1
OFLX gyrB I39V u	1	0	0	1
OFLX gyrB N538D u	1	0	0	1
OFLX gyrB N538Y u	1	0	0	1
OFLX gyrB P439R u	1	0	0	1
OFLX gyrB R485C u	1	0	0	1
OFLX gyrB T539N u	1	0	0	1
OFLX gyrB V496L u	1	0	0	1
OFLX gyrB V600A u	1	0	0	1

Para-aminosalicylic acid (PAS) Data, Abbreviations: counts = how many isolates had this mutation, no_data = how many isolates had this mutation but no data on whether they were phenotypically resistant or not to PAS, susceptible = how many isolates had this mutation but were phenotypically susceptible to PAS, resistant = how many isolates had this mutation but were resistant to PAS

country	PAS_SNP_CN_3073852_TC_thyA_H207R	PAS_SNP_CN_3074182_TC_thyA_Q97R	PAS_SNP_CN_3074449_AT_thyA_L8Q	PAS_inter-thyX-hsdS.1_G228A_u	PAS_thersnp	PAS_thyA_T202A_u	PAS_unknown
Azerbaijan	0	0	0	0	0	0	0
Bangladesh	0	0	0	0	0	0	0
Belarus	0	0	0	0	0	0	0
Brazil	0	0	0	0	0	0	0
Burma	0	0	0	0	0	0	0
Canada	0	0	0	0	0	0	0
China	1	0	0	0	0	0	0
Colombia	0	0	0	0	0	0	0
Democratic Republic of the Congo	0	0	0	0	0	0	0
Denmark	0	0	0	0	0	0	0
Djibouti	0	0	0	0	0	0	0
Dominican Republic	0	0	0	0	0	0	0
Estonia	0	0	0	0	0	0	0
Georgia	0	0	0	0	0	0	0
Germany	0	0	0	0	0	0	0
Guinea	0	0	0	0	0	0	0
India	0	0	0	0	0	0	0
Indonesia	0	0	0	0	0	0	0
Iran	0	0	0	0	0	0	0
Kazakhstan	0	0	0	0	0	0	0
Malawi	0	0	0	0	0	0	0
Mali	0	0	0	0	0	0	0
Moldova	0	0	0	0	0	0	0
Morocco	0	0	0	0	0	0	0
Nepal	0	0	0	0	0	0	0
Netherlands	0	0	0	0	0	0	0
Nigeria	0	0	0	0	0	0	0
Not Provided	5	2	4	3	15	18	5
Pakistan	0	0	0	0	0	0	0
Peru	10	0	4	0	1	0	0
Philippines	0	0	0	0	0	0	0
Portugal	0	0	0	0	0	0	0
Romania	0	0	0	0	0	0	0
Russia	0	1	0	0	1	0	0
Rwanda	0	0	0	0	0	0	0
Sierra Leone	0	0	0	0	0	0	0
South Africa	0	0	0	1	2	0	3
South Korea	0	0	0	6	6	0	3
Spain	0	0	0	0	0	0	0
Swaziland	0	0	0	0	0	0	0
Thailand	0	0	0	0	0	0	0
Turkmenistan	0	0	0	0	0	0	0
Uganda	0	0	0	0	0	0	0
United Kingdom	0	1	0	0	0	0	0
Uzbekistan	0	0	0	0	0	0	0
Vietnam	0	0	0	0	0	0	0

resistance mutation	counts	no data	susceptible	resistant
PAS thyA T202A u	18	0	0	18
PAS SNP CN 3073852 TC thyA H207R	16	14	0	2
PAS unknown	11	0	0	11
PAS upstream intergenic-thyX-hsdS.1 G228A u	10	0	0	10
PAS SNP CN 3074449 AT thyA L8Q	8	7	0	1
PAS SNP CN 3074182 TC thyA Q97R	4	3	0	1
PAS folC R49Q u	2	0	0	2
PAS folC R49W u	2	0	0	2
PAS upstream intergenic-thyX-hsdS.1 C235T u	2	0	0	2
PAS upstream intergenic-thyA G24A u	2	0	0	2
PAS folC E153G u	1	0	0	1
PAS folC I43S u	1	0	0	1
PAS folC I43T u	1	0	0	1
PAS folC I43V u	1	0	0	1
PAS folC S98G u	1	0	0	1
PAS upstream intergenic-thyX-hsdS.1 C226A u	1	0	0	1
PAS upstream intergenic-thyX-hsdS.1 G240A u	1	0	0	1
PAS upstream intergenic-thyA T117C u	1	0	0	1
PAS upstream intergenic-thyA d49CCGCAGCGACTCGCCGCCAAACAAACCCAGCGGGCGATCGCAAGCGCGGCGAAGCCG u	1	0	0	1
PAS thyA A262V u	1	0	0	1
PAS thyA D6G u	1	0	0	1
PAS thyA D81A u	1	0	0	1
PAS thyA E137G u	1	0	0	1
PAS thyA E58V u	1	0	0	1
PAS thyA F152V u	1	0	0	1
PAS thyA P145L u	1	0	0	1
PAS thyA T22P u	1	0	0	1
PAS SNP P 3074479 AG.157 thyA	0	0	0	0

Pyrazinamide (PZA) Data, Abbreviations: counts = how many isolates had this mutation, no_data = how many isolates had this mutation but no data on whether they were phenotypically resistant or not to PZA, susceptible = how many isolates had this mutation but were phenotypically susceptible to PZA, resistant = how many isolates had this mutation but were resistant to PZA

country	PZA_SNP_CN_2289090_TC_pncA_H51R	PZA_SNP_CN_2289099_TG_pncA_K48T	PZA_SNP_CN_2289213_TC_pncA_Q10R	PZA_SNP_CN_2289213_TG_pncA_Q10P	PZA_othersnp	PZA_promoter-pncA_T11C_u	PZA_unknown
Azerbaijan	0	0	0	1	0	0	0
Bangladesh	0	0	0	0	0	0	0
Belarus	0	0	2	0	56	0	0
Brazil	0	0	0	0	0	0	0
Burma	0	0	0	0	1	0	0
Canada	0	0	0	0	0	0	0
China	0	0	0	2	22	0	0
Colombia	0	0	0	0	0	0	0
Democratic Republic of the Congo	0	0	0	0	0	0	0
Denmark	0	0	0	0	0	0	0
Djibouti	0	0	0	0	0	0	0
Dominican Republic	0	0	0	0	0	0	0
Estonia	0	0	0	0	0	0	0
Georgia	0	0	0	0	0	0	0
Germany	0	0	0	0	14	0	5
Guinea	0	0	0	0	1	0	0
India	0	0	0	0	2	0	0
Indonesia	0	0	0	0	0	0	0
Iran	0	0	0	0	4	0	1
Kazakhstan	0	0	0	0	0	0	0
Malawi	0	0	0	0	4	0	0
Mali	0	0	0	0	3	0	0
Moldova	0	0	0	2	6	0	0
Morocco	0	0	0	0	0	0	0
Nepal	0	0	0	0	0	0	0
Netherlands	0	0	0	0	0	0	0
Nigeria	0	0	0	0	0	0	0
Not Provided	10	12	6	6	260	15	78
Pakistan	0	0	0	0	0	0	0
Peru	79	35	45	9	296	8	150
Philippines	0	0	0	0	0	0	0
Portugal	0	0	0	0	0	0	0
Romania	0	0	0	0	20	0	0
Russia	2	0	0	1	121	6	71
Rwanda	0	0	0	0	0	0	1
Sierra Leone	0	0	0	0	4	1	3
South Africa	0	0	0	3	112	0	66
South Korea	0	0	0	0	23	0	24
Spain	0	0	0	0	0	0	0
Swaziland	0	0	0	0	0	0	0
Thailand	0	0	0	0	0	0	0
Turkmenistan	0	0	0	0	0	1	2
Uganda	0	0	0	0	12	1	10
United Kingdom	0	0	0	0	41	2	19
Uzbekistan	1	0	2	39	69	17	40
Vietnam	0	0	0	0	0	0	0

resistance mutation	99	100	susceptible	resistant
PZA unknown	470	0	0	470
PZA SNP CN 2289090 TC pncA H51R	92	12	9	71
PZA SNP CN 2289213 TG pncA Q10P	63	11	4	48
PZA SNP CN 2289213 TC pncA Q10R	55	4	5	46
PZA upstream intergenic-pncA T11C u	51	0	0	51
PZA SNP CN 2289099 TG pncA K48T	47	9	22	16
PZA SNP CN 2289202 AG pncA C14R	46	12	0	34
PZA SNP CN 2288844 AG pncA I133T	40	1	16	23
PZA SNP CN 2289016 TG pncA T76P	38	23	2	13
PZA SNP CN 2288883 AG pncA L120P	35	15	0	20
PZA SNP CN 2288839 TG pncA T135P	34	12	2	20
PZA SNP CN 2289040 AC pncA W68G	31	22	5	4
PZA SNP CN 2289073 GC pncA H57D	26	5	0	21
PZA SNP CN 2289231 AG pncA L4S	25	8	0	17
PZA SNP CN 2288826 AG pncA V139A	23	5	7	11
PZA SNP CN 2289142 AC pncA Y34D	22	19	1	2
PZA SNP CN 2288820 TG pncA Q141P	20	7	2	11
PZA SNP CN 2289081 GA pncA P54L	20	4	4	12
PZA SNP CN 2289180 AC pncA V21G	20	5	4	11
PZA pncA V125G u	20	0	0	20
PZA SNP CN 2289207 TG pncA D12A	19	6	4	9
PZA SNP CN 2288727 AG pncA L172P	17	11	2	4
PZA SNP CN 2289097 CT pncA D49N	16	3	1	12
PZA SNP CN 2288952 CT pncA G97D	15	9	0	6
PZA SNP CN 2288953 CT pncA G97S	15	3	1	11
PZA SNP CN 2289040 AG pncA W68R	15	4	2	9
PZA DEL F 2288939 d302TCCGGTGTAG pncA	13	9	0	4
PZA SNP CN 2288805 GA pncA A146V	13	9	0	4
PZA SNP CN 2289054 TG pncA D63A	12	2	3	7
PZA SNP CN 2289070 AG pncA F58L	12	4	3	5
PZA SNP CN 2288805 GT pncA A146E	11	0	4	7
PZA SNP CN 2289207 TC pncA D12G	11	1	2	8
PZA pncA I6L u	11	0	0	11
PZA DEL F 2289069 d172A pncA F58L	10	6	0	4
PZA SNP CN 2288703 AC pncA V180G	10	2	0	8
PZA SNP CN 2288764 TC pncA T160A	10	3	0	7
PZA SNP CN 2288827 CT pncA V139M	10	3	1	6
PZA SNP CN 2289220 CT pncA D8N	10	6	0	4
PZA SNP CN 2288704 CA pncA V180F	9	0	2	7
PZA INS F 2289009 i232C pncA G78G	8	3	1	4
PZA INS F 2289050 i191T pncA Y64	8	3	1	4
PZA SNP CN 2288850 AC pncA V131G	8	3	1	4
PZA SNP CN 2288859 AC pncA V128G	8	1	2	5
PZA upstream intergenic-pncA T11G u	8	0	0	8
PZA SNP CN 2288778 AC pncA V155G	7	3	1	3
PZA SNP CN 2288943 GA pncA T100I	7	0	1	6
PZA SNP CN 2288988 AG pncA L85P	7	2	1	4
PZA SNP CN 2289043 AG pncA S67P	7	1	0	6
PZA SNP CN 2289072 TC pncA H57R	7	2	0	5
PZA SNP CN 2288696 CA pncA L182F	6	0	5	1
PZA SNP CN 2288835 TC pncA D136G	6	3	1	2
PZA SNP CN 2288965 CA pncA V93L	6	0	1	5
PZA SNP CN 2289072 TA pncA H57L	6	0	1	5
PZA SNP CN 2289103 TC pncA T47A	6	2	3	1
PZA DEL F 2288923 d318C pncA	5	2	1	2
PZA INS F 2288825 i416C pncA	5	1	0	4
PZA INS F 2288851 i390CC pncA	5	1	1	3
PZA INS F 2288942 i299T pncA	5	2	1	2
PZA SNP CN 2288697 AG pncA L182S	5	2	2	1
PZA SNP CN 2288718 AG pncA M175T	5	2	0	3
PZA SNP CN 2288938 CG pncA A102P	5	2	2	1
PZA SNP CN 2288944 TG pncA T100P	5	1	0	4
PZA SNP CN 2288956 TC pncA K96E	5	1	0	4
PZA SNP CN 2288973 AG pncA I90T	5	1	1	3
PZA SNP CN 2289028 AG pncA C72R	5	2	1	2
PZA SNP CN 2289030 TC pncA H71R	5	0	3	2
PZA SNP CN 2289162 AG pncA L27P	5	1	2	2
PZA SNP CN 2289219 TC pncA D8G	5	2	1	2
PZA pncA C138R u	5	0	0	5
PZA pncA D8E u	5	0	0	5
PZA pncA S164P u	5	0	0	5
PZA INS F 2288851 i390C pncA	4	2	0	2
PZA INS F 2288887 i354A pncA	4	2	0	2
PZA SNP CN 2288766 AC pncA L159R	4	0	0	4
PZA SNP CN 2288818 TC pncA T142A	4	2	0	2
PZA SNP CN 2288841 GA pncA A134V	4	2	0	2
PZA SNP CN 2288955 TG pncA K96T	4	1	0	3
PZA SNP CN 2289037 GA pncA P69S	4	0	4	0
PZA pncA F13L u	4	0	0	4
PZA pncA M1T u	4	0	0	4
PZA pncA R140P u	4	0	0	4
PZA pncA V7G u	4	0	0	4
PZA DEL F 2288776 d465GCACCCTG pncA	3	1	1	1
PZA INS F 2288725 i516C pncA	3	1	0	2
PZA SNP CN 2288775 AG pncA L156P	3	2	0	1
PZA SNP CN 2288883 AC pncA L120R	3	1	0	2
PZA SNP CN 2289001 AC pncA F81V	3	0	2	1
PZA SNP CN 2289015 GA pncA T76I	3	1	1	1
PZA SNP CN 2289054 TC pncA D63G	3	0	1	2
PZA SNP CN 2289073 GA pncA H57Y	3	0	0	3
PZA pncA C14G u	3	0	0	3
PZA pncA D136Y u	3	0	0	3
PZA pncA H51Q u	3	0	0	3
PZA pncA H82R u	3	0	0	3
PZA pncA L116R u	3	0	0	3
PZA pncA L4W u	3	0	0	3
PZA pncA P54Q u	3	0	0	3
PZA pncA P62S u	3	0	0	3
PZA upstream intergenic-pncA A12G u	3	0	0	3

PZA SNP CN 2288748 AC pncA M173R	2	0	0	1	1
PZA SNP CN 2288742 GA pncA T167I	2	1	0	0	1
PZA SNP CN 2288818 TG pncA T142P	2	0	0	0	2
PZA SNP CN 2288826 AC pncA V139G	2	0	0	0	1
PZA SNP CN 2288869 CA pncA V125F	2	0	0	0	2
PZA SNP CN 2288887 AC pncA W119G	2	0	0	1	1
PZA SNP CN 2288919 CT pncA G108E	2	1	0	0	1
PZA SNP CN 2288925 AG pncA F106S	2	0	0	1	1
PZA SNP CN 2288964 AC pncA V93G	2	0	0	1	1
PZA SNP CN 2289009 CA pncA G78V	2	0	0	0	2
PZA SNP CN 2289030 TG pncA H71P	2	1	0	0	1
PZA SNP CN 2289046 AG pncA S66P	2	1	0	0	1
PZA SNP CN 2289090 TG pncA H51P	2	0	0	1	1
PZA SNP CN 2289091 GA pncA H51Y	2	2	0	0	0
PZA SNP CN 2289095 GC pncA D49E	2	1	0	0	1
PZA SNP CN 2289212 CG pncA Q10H	2	0	0	0	2
PZA SNP CN 2289216 AC pncA V9G	2	1	0	0	1
PZA SNP CN 2289228 AG pncA I5T	2	0	0	0	2
PZA pncA A46V u	2	0	0	0	2
PZA pncA D63H u	2	0	0	0	2
PZA pncA D8A u	2	0	0	0	2
PZA pncA F94L u	2	0	0	0	2
PZA pncA G97C u	2	0	0	0	2
PZA pncA H71Q u	2	0	0	0	2
PZA pncA H71Y u	2	0	0	0	2
PZA pncA H82D u	2	0	0	0	2
PZA pncA I6M u	2	0	0	0	2
PZA pncA K48E u	2	0	0	0	2
PZA pncA L35R u	2	0	0	0	2
PZA pncA M175V u	2	0	0	0	2
PZA pncA P62L u	2	0	0	0	2
PZA pncA P69R u	2	0	0	0	2
PZA pncA S104R u	2	0	0	0	2
PZA pncA V130M u	2	0	0	0	2
PZA pncA V155M u	2	0	0	0	2
PZA pncA V44G u	2	0	0	0	2
PZA pncA Y103C u	2	0	0	0	2
PZA pncA Y95S u	2	0	0	0	2
PZA upstream intergenic-pncA A7G u	2	0	0	0	2
PZA INS F 2288835 1406T pncA	1	0	0	0	1
PZA SNP CN 2288697 AC pncA L182W	1	1	0	0	0
PZA SNP CN 2288730 GA pncA A171V	1	1	0	0	0
PZA SNP CN 2288817 GA pncA T142M	1	1	0	0	0
PZA SNP CN 2288847 CT pncA G132D	1	0	0	0	1
PZA SNP CN 2288935 AG pncA Y103H	1	0	0	0	1
PZA SNP CN 2289042 GC pncA S67W	1	0	0	0	1
PZA SNP CN 2289069 AC pncA F58C	1	0	0	0	1
PZA SNP CN 2289186 AG pncA L19P	1	0	0	0	1
PZA pncA A102T u	1	0	0	0	1
PZA pncA A146P u	1	0	0	0	1
PZA pncA A171E u	1	0	0	0	1
PZA pncA A171T u	1	0	0	0	1
PZA pncA A3E u	1	0	0	0	1
PZA pncA A46E u	1	0	0	0	1
PZA pncA A46P u	1	0	0	0	1
PZA pncA C14W u	1	0	0	0	1
PZA pncA D129Y u	1	0	0	0	1
PZA pncA D12N u	1	0	0	0	1
PZA pncA D136N u	1	0	0	0	1
PZA pncA D49G u	1	0	0	0	1
PZA pncA F131 u	1	0	0	0	1
PZA pncA F58S u	1	0	0	0	1
PZA pncA F81C u	1	0	0	0	1
PZA pncA G105D u	1	0	0	0	1
PZA pncA G108R u	1	0	0	0	1
PZA pncA G132A u	1	0	0	0	1
PZA pncA G162D u	1	0	0	0	1
PZA pncA G17D u	1	0	0	0	1
PZA pncA G24D u	1	0	0	0	1
PZA pncA G78C u	1	0	0	0	1
PZA pncA G78S u	1	0	0	0	1
PZA pncA G97R u	1	0	0	0	1
PZA pncA H43P u	1	0	0	0	1
PZA pncA H133S u	1	0	0	0	1
PZA pncA I6T u	1	0	0	0	1
PZA pncA K96R u	1	0	0	0	1
PZA pncA L151S u	1	0	0	0	1
PZA pncA L156Q u	1	0	0	0	1
PZA pncA L159V u	1	0	0	0	1
PZA pncA L85R u	1	0	0	0	1
PZA pncA M175I u	1	0	0	0	1
PZA pncA M175K u	1	0	0	0	1
PZA pncA M1I u	1	0	0	0	1
PZA pncA P62R u	1	0	0	0	1
PZA pncA R154G u	1	0	0	0	1
PZA pncA R2W u	1	0	0	0	1
PZA pncA S104G u	1	0	0	0	1
PZA pncA S32I u	1	0	0	0	1
PZA pncA T100A u	1	0	0	0	1
PZA pncA T114P u	1	0	0	0	1
PZA pncA T160K u	1	0	0	0	1
PZA pncA T160P u	1	0	0	0	1
PZA pncA T168P u	1	0	0	0	1
PZA pncA V139L u	1	0	0	0	1
PZA pncA V155A u	1	0	0	0	1
PZA pncA V163A u	1	0	0	0	1
PZA pncA V7F u	1	0	0	0	1
PZA pncA V7L u	1	0	0	0	1
PZA pncA W119R u	1	0	0	0	1
PZA pncA W68C u	1	0	0	0	1
PZA pncA W68L u	1	0	0	0	1
PZA pncA Y103D u	1	0	0	0	1
PZA pncA Y64D u	1	0	0	0	1
PZA upstream intergenic-pncA T11A u	1	0	0	0	1
PZA upstream intergenic-pncA T15C u	1	0	0	0	1
PZA upstream intergenic-pncA T16C u	1	0	0	0	1
PZA upstream intergenic-pncA d10A u	1	0	0	0	1
PZA upstream intergenic-pncA d14T u	1	0	0	0	1
PZA DEL F 2288697 d544AACT pncA	0	0	0	0	0
PZA DEL F 2289060 d181GTGCCGGA pncA	0	0	0	0	0
PZA DEL N 2288942 d299GGTGA pncA	0	0	0	0	0
PZA SNP CN 2288784 GT pncA T153N	0	0	0	0	0
PZA SNP CN 2288848 CT pncA G132S	0	0	0	0	0
PZA SNP CN 2288853 AC pncA V130G	0	0	0	0	0
PZA SNP CN 2288853 AT pncA V130E	0	0	0	0	0
PZA SNP CN 2288878 GA pncA Q122	0	0	0	0	0
PZA SNP CN 2288933 GC pncA Y103	0	0	0	0	0
PZA SNP CN 2288956 TG pncA K96Q	0	0	0	0	0
PZA SNP CN 2289042 GT pncA S67	0	0	0	0	0
PZA SNP CN 2289050 AT pncA Y64	0	0	0	0	0
PZA SNP CN 2289150 AC pncA I31S	0	0	0	0	0
PZA SNP CN 2289206 GC pncA D12E	0	0	0	0	0
PZA SNP CN 2289214 GA pncA Q10	0	0	0	0	0
PZA SNP P 2289245 TA.37 pncA	0	0	0	0	0
PZA SNP P 2289251 AC.31 pncA	0	0	0	0	0
PZA SNP P 2289252 TC.30 pncA	0	0	0	0	0
PZA SNP P 2289252 TG.30 pncA	0	0	0	0	0

Rifamycin (RIF) Data, Abbreviations: counts = how many isolates had this mutation, no_data = how many isolates had this mutation but no data on whether they were phenotypically resistant or not to RIF, susceptible = how many isolates had this mutation but were phenotypically susceptible to RIF, resistant = how many isolates had this mutation but were resistant to RIF

country	RIF_SNP_CN_761109_GT_rpoB_D435Y	RIF_SNP_CN_761110_AT_rpoB_D435V	RIF_SNP_CN_761139_CT_rpoB_H445Y	RIF_SNP_CN_761155_CT_rpoB_S450L	RIF_SNP_CN_761161_TC_rpoB_L452P	RIF_othersnp	RIF_unknown
Azerbaijan	0	0	0	1	0	0	0
Bangladesh	0	0	0	0	0	0	1
Belarus	3	0	1	77	0	20	0
Brazil	0	0	0	0	0	0	0
Burma	0	0	0	1	0	0	0
Canada	0	0	0	0	0	0	0
China	3	4	11	57	9	36	5
Colombia	0	0	0	0	0	0	1
Democratic Republic of the Congo	0	0	0	0	0	0	0
Denmark	0	0	0	0	0	0	0
Djibouti	0	0	0	0	0	0	0
Dominican Republic	0	0	0	0	0	0	0
Estonia	0	0	0	2	0	0	0
Georgia	0	0	0	4	1	0	0
Germany	0	1	1	13	1	5	1
Guinea	0	0	0	1	0	0	0
India	0	2	0	8	0	0	0
Indonesia	0	0	0	0	0	0	0
Iran	0	0	0	10	0	2	1
Kazakhstan	0	0	0	2	0	0	0
Malawi	0	1	0	4	0	1	1
Mali	0	8	0	4	3	5	0
Moldova	1	0	0	34	0	4	0
Morocco	0	0	0	0	0	1	0
Nepal	0	0	0	1	0	1	0
Netherlands	0	0	0	0	0	0	0
Nigeria	0	0	0	1	0	0	0
Not Provided	42	73	49	515	18	123	57
Pakistan	0	0	0	1	0	0	0
Peru	22	134	28	491	10	110	57
Philippines	0	0	0	1	0	0	1
Portugal	0	0	0	1	0	0	0
Romania	6	4	2	19	0	3	0
Russia	6	7	3	350	6	28	74
Rwanda	0	1	0	3	0	0	1
Sierra Leone	2	0	3	6	1	2	1
South Africa	10	55	15	188	62	36	8
South Korea	4	3	2	35	1	1	3
Spain	0	0	0	1	1	0	0
Swaziland	0	0	0	2	0	0	0
Thailand	0	0	0	0	0	0	0
Turkmenistan	0	0	0	1	0	0	0
Uganda	1	3	2	25	0	14	4
United Kingdom	3	5	7	54	6	65	3
Uzbekistan	3	8	7	213	4	18	13
Vietnam	0	0	0	0	0	0	0

resistance mutation	counts	no_data	susceptible	resistant
RIF SNP CN 761155 CT rpoB S450L	2126	232	228	1666
RIF SNP CN 761110 AT rpoB D435V	309	19	27	263
RIF unknown	232	0	0	232
RIF SNP CN 761139 CT rpoB H445Y	131	16	10	105
RIF SNP CN 761161 TC rpoB L452P	123	5	75	43
RIF SNP CN 761109 GT rpoB D435Y	106	18	22	66
RIF SNP CN 761139 CG rpoB H445D	99	21	10	68
RIF SNP CN 761140 AG rpoB H445R	59	9	3	47
RIF SNP CN 761155 CG rpoB S450W	50	3	15	32
RIF SNP CN 761140 AT rpoB H445L	40	7	4	29
RIF SNP CN 761095 TC rpoB L430P	37	3	20	14
RIF SNP CN 761140 AT rpoB I491F	36	8	18	10
RIF SNP CN 761277 AT rpoB I491F	36	8	18	10
RIF SNP CN 760314 GT rpoB V170F	29	13	1	15
RIF rpoB S450P u	9	0	0	9
RIF rpoB Q432P u	7	0	0	7
RIF rpoB S441L u	7	0	0	7
RIF rpoB D435G u	6	0	0	6
RIF rpoB Q432K u	6	0	0	6
RIF rpoB R448Q u	6	0	0	6
RIF rpoB E250G u	4	0	0	4
RIF rpoB H445N u	4	0	0	4
RIF rpoB V695L u	4	0	0	4
RIF rpoB P454L u	3	0	0	3
RIF rpoB Q432L u	3	0	0	3
RIF rpoB D435N u	2	0	0	2
RIF rpoB H445P u	2	0	0	2
RIF rpoB L494P u	2	0	0	2
RIF rpoB M434I u	2	0	0	2
RIF rpoB Q432E u	2	0	0	2
RIF rpoB A451G u	1	0	0	1
RIF rpoB D435A u	1	0	0	1
RIF rpoB G675D u	1	0	0	1
RIF rpoB G981D u	1	0	0	1
RIF rpoB H835Q u	1	0	0	1
RIF rpoB H835R u	1	0	0	1
RIF rpoB I925V u	1	0	0	1
RIF rpoB K446Q u	1	0	0	1
RIF rpoB L430R u	1	0	0	1
RIF rpoB L443W u	1	0	0	1
RIF rpoB M707T u	1	0	0	1
RIF rpoB P280L u	1	0	0	1
RIF rpoB P439S u	1	0	0	1
RIF rpoB S428R u	1	0	0	1
RIF rpoB S441T u	1	0	0	1
RIF rpoB T400A u	1	0	0	1
RIF rpoB T400P u	1	0	0	1
RIF rpoB T444P u	1	0	0	1
RIF rpoB V113I u	1	0	0	1
RIF rpoB V469L u	1	0	0	1
RIF SNP CN 761102 AC rpoB Q432H	0	0	0	0
RIF SNP CN 761155 CA rpoB S450.	0	0	0	0

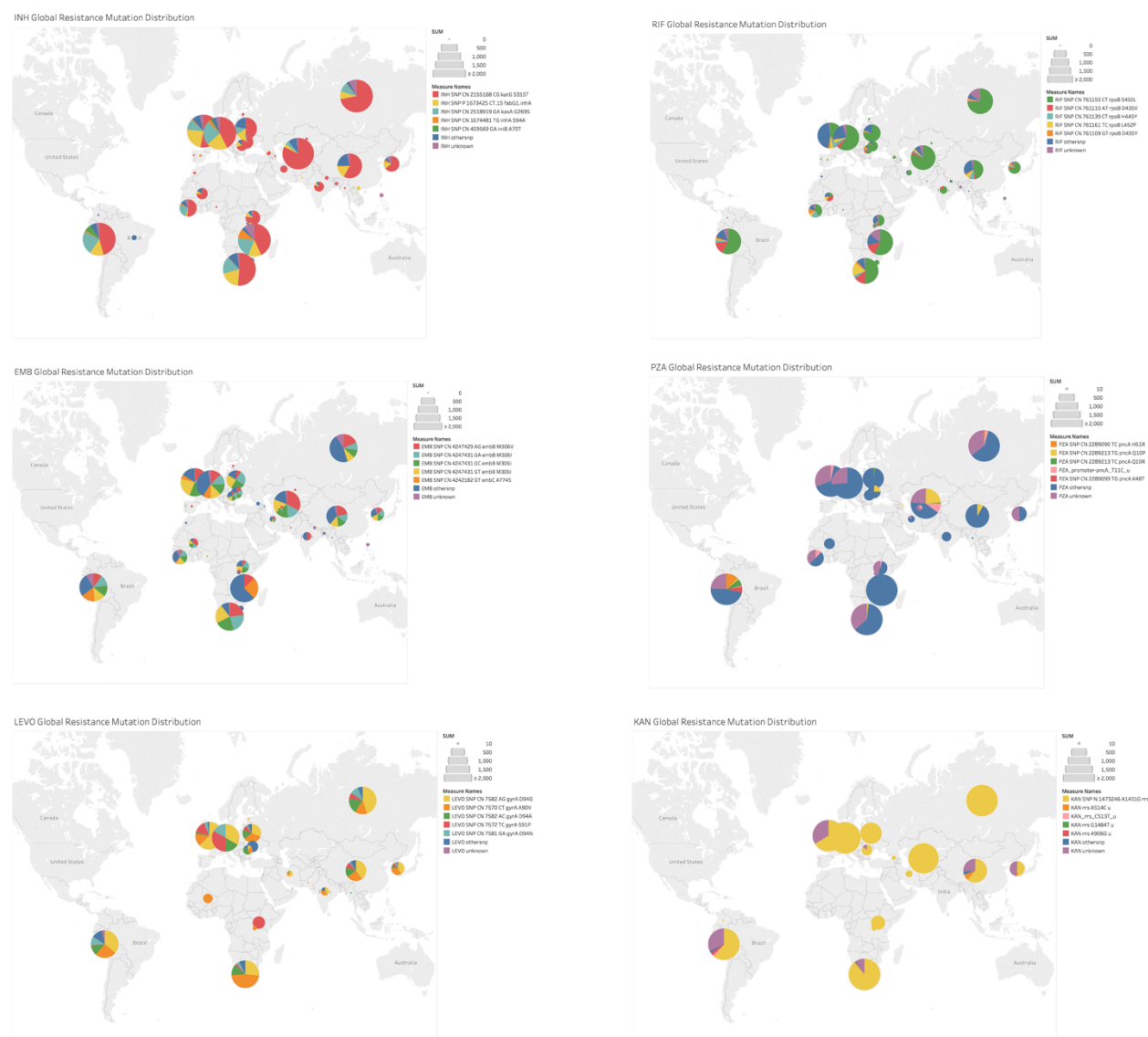
Streptomycin (STR) Data, Abbreviations: counts = how many isolates had this mutation, no_data = how many isolates had this mutation but no data on whether they were phenotypically resistant or not to RIF, susceptible = how many isolates had this mutation but were phenotypically susceptible to STR, resistant = how many isolates had this mutation but were resistant to STR

country	STR_SNP_CN_4407967_AG_gid_L79S	STR_SNP_CN_781687_AG_rpsL_K43R	STR_SNP_CN_781822_AG_rpsL_K88R	STR_SNP_N_1472359_A514C_rrs	STR_SNP_N_1472362_C517T_rrs	STR_othersnp	STR_unknown
Azerbaijan	0	0	0	0	1	0	0
Bangladesh	0	0	1	0	0	0	1
Belarus	0	63	6	20	10	1	0
Brazil	0	1	0	0	0	0	1
Burma	0	1	0	0	0	0	0
Canada	0	0	0	0	0	0	0
China	0	46	16	10	2	4	7
Colombia	0	0	0	0	0	0	0
Democratic Republic of the Congo	0	0	0	0	0	0	0
Denmark	0	0	0	0	0	0	0
Djibouti	0	0	0	0	0	0	0
Dominican Republic	0	0	0	0	0	0	0
Estonia	0	2	0	0	0	0	0
Georgia	0	5	0	0	0	0	0
Germany	3	29	3	3	2	38	11
Guinea	0	0	0	0	1	0	0
India	0	9	1	0	1	2	2
Indonesia	0	0	0	0	0	0	0
Iran	0	8	0	1	3	0	1
Kazakhstan	0	2	0	0	0	0	0
Malawi	0	4	6	0	0	18	3
Mali	17	0	1	1	0	1	1
Moldova	0	13	27	0	1	2	0
Morocco	0	0	0	0	0	0	1
Nepal	0	2	0	1	0	0	0
Netherlands	0	9	1	1	7	0	0
Nigeria	0	0	0	0	0	0	1
Not Provided	15	239	29	39	20	244	50
Pakistan	0	0	0	0	1	0	0
Peru	2	71	8	14	7	340	96
Philippines	0	0	0	0	0	0	1
Portugal	0	1	0	0	0	0	0
Romania	0	9	2	2	0	6	5
Russia	0	221	43	17	184	19	47
Rwanda	0	0	0	0	0	0	1
Sierra Leone	0	9	7	0	0	13	4
South Africa	74	72	15	83	7	9	15
South Korea	0	12	1	2	0	3	3
Spain	0	0	0	0	0	0	1
Swaziland	0	0	1	0	0	0	0
Thailand	0	0	0	0	0	0	0
Turkmenistan	0	5	1	0	0	0	0
Uganda	0	2	3	0	2	13	9
United Kingdom	5	53	10	16	5	55	5
Uzbekistan	0	212	12	13	3	13	5
Vietnam	0	2	0	0	0	0	0

resistance mutation	counts	no_data	susceptible	resistant
STR SNP CN 781687 AG rpsL K43R	1102	172	112	818
STR unknown	271	0	0	271
STR SNP N 1472362 C517T rrs	257	14	54	189
STR SNP N 1472359 A514C rrs	223	54	25	144
STR SNP CN 781822 AG rpsL K88R	194	22	44	128
STR SNP CN 4407967 AG gid L79S	116	27	13	76
STR SNP CN 4407967 AC gid L79W	73	15	26	32
STR SNP CN 781822 AC rpsL K88T	55	8	6	41
STR SNP CN 4407934 AC gid L90R	35	6	10	19
STR SNP CN 4407768 CA gid L145F	34	3	22	9
STR SNP N 1472358 C513T rrs	34	6	2	26
STR SNP N 1472751 A906G rrs	29	3	8	18
STR DEL F 4407640 d562A gid	28	0	17	11
STR SNP CN 4408094 CT gid G37R	28	0	17	11
STR DEL F 4407852 d350C gid	27	19	5	3
STR SNP CN 4407995 TG gid S70R	27	19	5	3
STR SNP CN 4407985 CG gid G73A	23	0	8	15
STR SNP CN 4407809 CA gid D132Y	21	5	1	15
STR SNP CN 4407916 CA gid R96L	16	3	5	8
STR gid L50R u	14	0	0	14
STR SNP N 1472753 A908C rrs	11	0	3	8
STR SNP CN 4407748 AG gid L152S	10	1	5	4
STR gid P84L u	10	0	0	10
STR SNP CN 4408064 GA gid R47W	9	2	5	2
STR rrs A514T u	8	0	0	8
STR SNP CN 4407832 AG gid V124A	7	0	2	5
STR SNP CN 781822 AT rpsL K88M	7	1	2	4
STR gid G76C u	7	0	0	7
STR gid A138V u	6	0	0	6
STR gid D67G u	6	0	0	6
STR gid L26F u	6	0	0	6
STR SNP CN 4408148 CG gid A19P	5	0	2	3
STR gid H48Y u	5	0	0	5
STR gid P75R u	5	0	0	5
STR gid R137P u	5	0	0	5
STR gid V77G u	5	0	0	5
STR SNP CN 4408102 CT gid G34E	4	2	0	2
STR gid A10P u	4	0	0	4
STR gid S149R u	4	0	0	4
STR SNP CN 4407947 GA gid L86F	3	0	3	0
STR SNP CN 4408060 TG gid H48P	3	1	1	1
STR gid A134E u	3	0	0	3
STR gid A138E u	3	0	0	3
STR gid A80P u	3	0	0	3
STR gid D85A u	3	0	0	3
STR gid E92Q u	3	0	0	3
STR gid G69D u	3	0	0	3
STR gid L86P u	3	0	0	3
STR gid P75S u	3	0	0	3
STR gid R118S u	3	0	0	3
STR gid T146K u	3	0	0	3
STR gid V77A u	3	0	0	3
STR rrs C905A u	3	0	0	3
STR SNP N 1473167 T1322G rrs	2	0	0	2
STR gid A200E u	2	0	0	2
STR gid E170G u	2	0	0	2
STR gid E40A u	2	0	0	2
STR gid G117E u	2	0	0	2
STR gid G117V u	2	0	0	2
STR gid G30R u	2	0	0	2
STR gid G34A u	2	0	0	2
STR gid G34V u	2	0	0	2
STR gid G37E u	2	0	0	2
STR gid H48N u	2	0	0	2
STR gid L101F u	2	0	0	2
STR gid L108R u	2	0	0	2
STR gid L18H u	2	0	0	2
STR gid L59R u	2	0	0	2
STR gid P75L u	2	0	0	2
STR gid R137W u	2	0	0	2
STR gid R83G u	2	0	0	2
STR gid R83W u	2	0	0	2
STR gid V110A u	2	0	0	2
STR gid V65G u	2	0	0	2
STR gid W45S u	2	0	0	2
STR rpsL K88Q u	2	0	0	2
STR SNP CN 4408138 TC gid Y22C	1	0	0	1
STR gid A10V u	1	0	0	1
STR gid A119D u	1	0	0	1
STR gid A133P u	1	0	0	1
STR gid A138P u	1	0	0	1
STR gid A140S u	1	0	0	1
STR gid A19G u	1	0	0	1
STR gid A205E u	1	0	0	1
STR gid A205T u	1	0	0	1
STR gid A27P u	1	0	0	1
STR gid A72V u	1	0	0	1
STR gid A82E u	1	0	0	1
STR gid C52F u	1	0	0	1
STR gid D67A u	1	0	0	1
STR gid E120K u	1	0	0	1
STR gid E121D u	1	0	0	1
STR gid G117R u	1	0	0	1
STR gid G117W u	1	0	0	1
STR gid G130A u	1	0	0	1
STR gid G182R u	1	0	0	1
STR gid G30D u	1	0	0	1
STR gid G30V u	1	0	0	1
STR gid G34R u	1	0	0	1

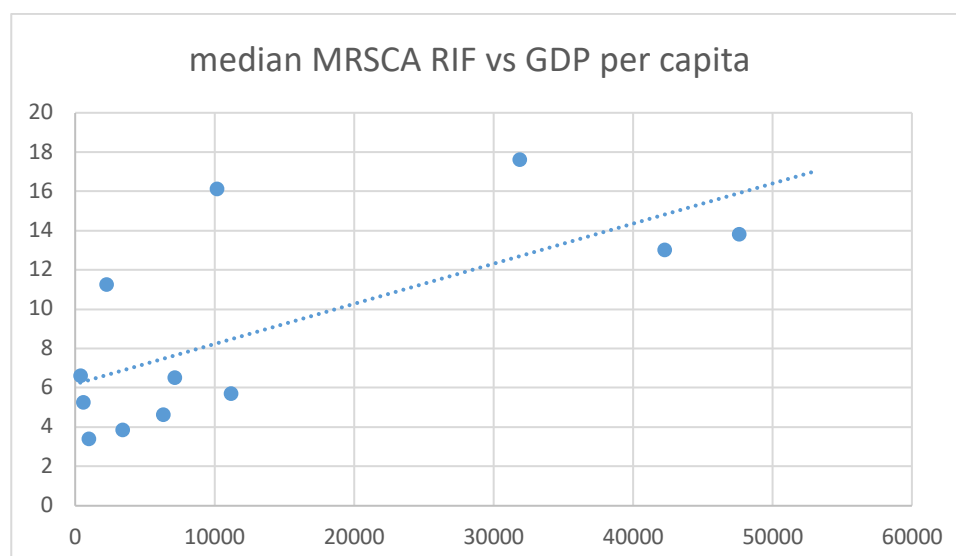
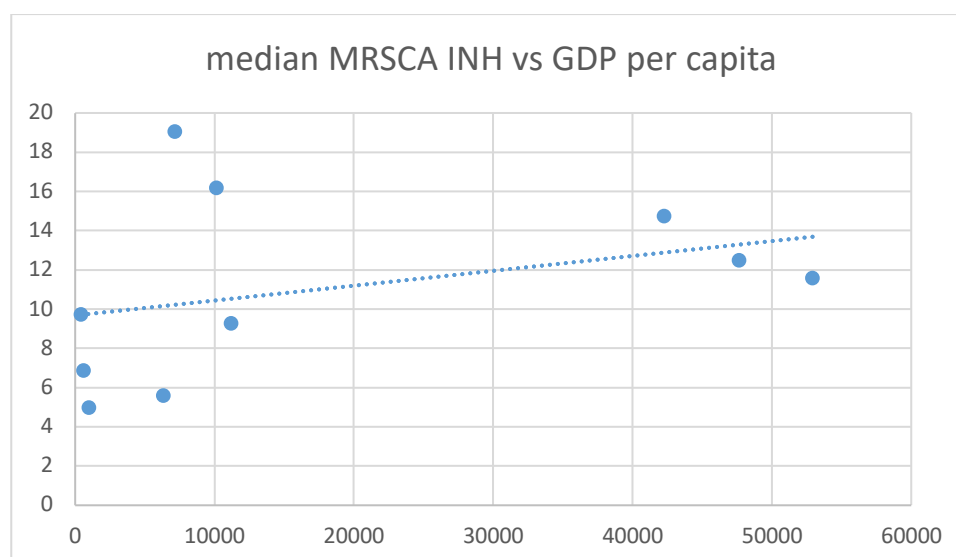
STR gid G34W u	1	0	0	1
STR gid G71R u	1	0	0	1
STR gid H48D u	1	0	0	1
STR gid H48Q u	1	0	0	1
STR gid H48R u	1	0	0	1
STR gid I11N u	1	0	0	1
STR gid I162S u	1	0	0	1
STR gid I81R u	1	0	0	1
STR gid I81T u	1	0	0	1
STR gid L108P u	1	0	0	1
STR gid L142W u	1	0	0	1
STR gid L160P u	1	0	0	1
STR gid L79F u	1	0	0	1
STR gid L86R u	1	0	0	1
STR gid L91P u	1	0	0	1
STR gid M150R u	1	0	0	1
STR gid P78L u	1	0	0	1
STR gid P93L u	1	0	0	1
STR gid P93Q u	1	0	0	1
STR gid P93R u	1	0	0	1
STR gid R116W u	1	0	0	1
STR gid R118C u	1	0	0	1
STR gid R118L u	1	0	0	1
STR gid R137L u	1	0	0	1
STR gid R175P u	1	0	0	1
STR gid R20P u	1	0	0	1
STR gid R21W u	1	0	0	1
STR gid R33P u	1	0	0	1
STR gid R47G u	1	0	0	1
STR gid R47P u	1	0	0	1
STR gid R64W u	1	0	0	1
STR gid R83P u	1	0	0	1
STR gid R96C u	1	0	0	1
STR gid S131T u	1	0	0	1
STR gid S70N u	1	0	0	1
STR gid T98I u	1	0	0	1
STR gid V105E u	1	0	0	1
STR gid V115G u	1	0	0	1
STR gid V203L u	1	0	0	1
STR gid V411 u	1	0	0	1
STR gid V66A u	1	0	0	1
STR gid V88A u	1	0	0	1
STR gid V89G u	1	0	0	1
STR gid Y195C u	1	0	0	1
STR gid Y22S u	1	0	0	1
STR rpsL E70A u	1	0	0	1
STR rpsL E76A u	1	0	0	1
STR rrs A1012G u	1	0	0	1
STR rrs A1223G u	1	0	0	1
STR rrs A1325C u	1	0	0	1
STR rrs A170G u	1	0	0	1
STR rrs A504C u	1	0	0	1
STR rrs A700T u	1	0	0	1
STR rrs A703G u	1	0	0	1
STR rrs A753T u	1	0	0	1
STR rrs A907C u	1	0	0	1
STR rrs A907T u	1	0	0	1
STR rrs A908G u	1	0	0	1
STR rrs A948T u	1	0	0	1
STR rrs C1067T u	1	0	0	1
STR rrs C1199G u	1	0	0	1
STR rrs C1390T u	1	0	0	1
STR rrs C1402T u	1	0	0	1
STR rrs C239A u	1	0	0	1
STR rrs C244T u	1	0	0	1
STR rrs C270T u	1	0	0	1
STR rrs C332T u	1	0	0	1
STR rrs C397T u	1	0	0	1
STR rrs C414A u	1	0	0	1
STR rrs C499T u	1	0	0	1
STR rrs C546T u	1	0	0	1
STR rrs C692T u	1	0	0	1
STR rrs C708T u	1	0	0	1
STR rrs C897T u	1	0	0	1
STR rrs C905G u	1	0	0	1
STR rrs C936T u	1	0	0	1
STR rrs G1042C u	1	0	0	1
STR rrs G1072A u	1	0	0	1
STR rrs G1415T u	1	0	0	1
STR rrs G319A u	1	0	0	1
STR rrs G395C u	1	0	0	1
STR rrs G407A u	1	0	0	1
STR rrs G408T u	1	0	0	1
STR rrs G537A u	1	0	0	1
STR rrs G544A u	1	0	0	1
STR rrs G685A u	1	0	0	1
STR rrs G749C u	1	0	0	1
STR rrs G771A u	1	0	0	1
STR rrs G878A u	1	0	0	1
STR rrs G887T u	1	0	0	1
STR rrs G909T u	1	0	0	1
STR rrs T1206C u	1	0	0	1
STR rrs T1208A u	1	0	0	1
STR rrs T1217A u	1	0	0	1
STR rrs T16C u	1	0	0	1
STR rrs T327C u	1	0	0	1
STR rrs T411A u	1	0	0	1
STR rrs T529G u	1	0	0	1
STR rrs T545A u	1	0	0	1
STR rrs T580C u	1	0	0	1
STR rrs T672A u	1	0	0	1
STR rrs T696G u	1	0	0	1
STR DEL F 4408023 d179T gid	0	0	0	0
STR DEL F 4408116 d86G gid	0	0	0	0
STR SNP CN 4408091 GT gid P38T	0	0	0	0
STR SNP I 1473637 A.21 rrs.rfl	0	0	0	0
STR SNP N 1473109 T1264G rrs	0	0	0	0
STR SNP N 1473343 G1498T rrs	0	0	0	0

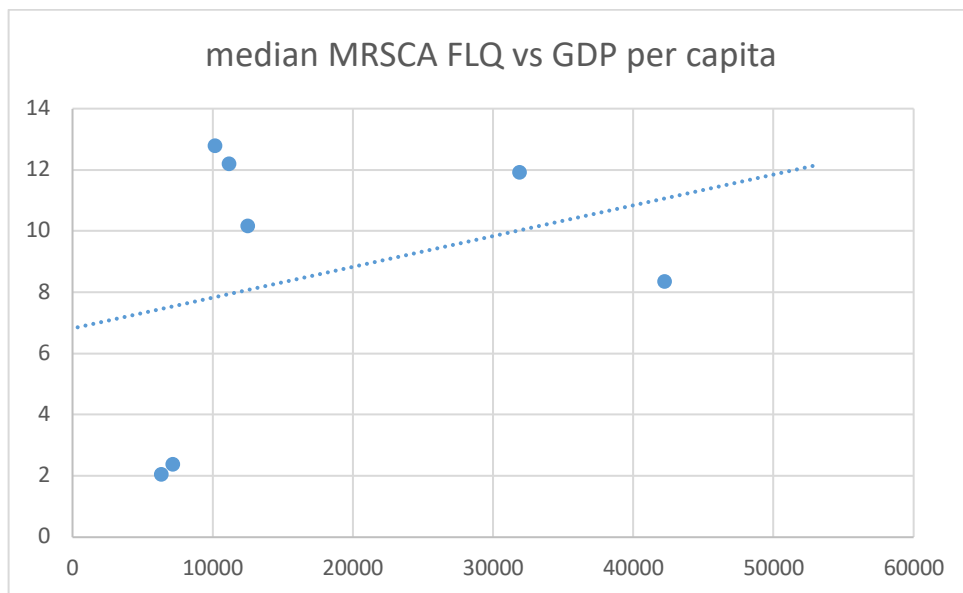
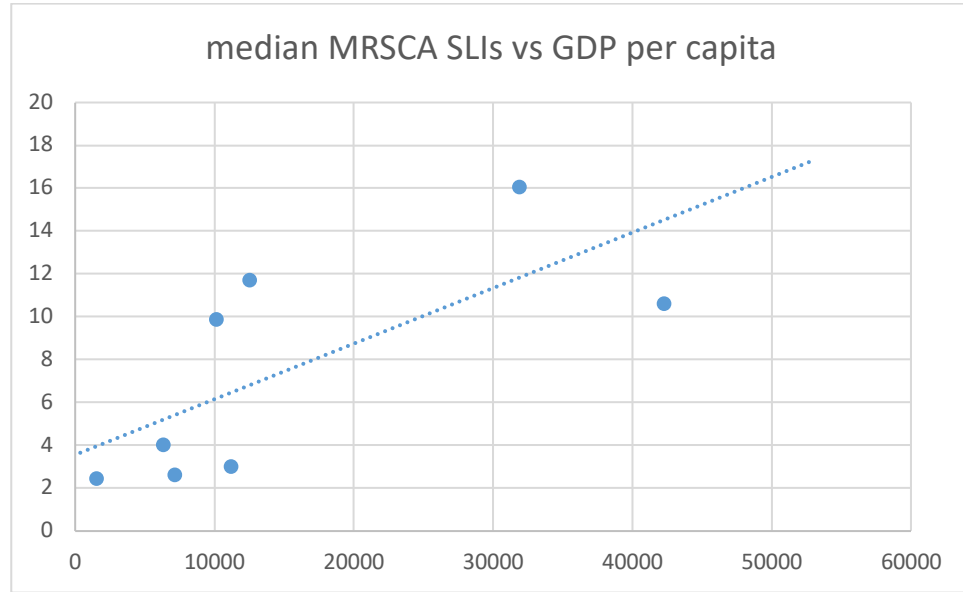
Supplementary Figure 7 Distribution of resistance mutations for six drugs (n=9385). Pie chart size is proportional isolate number from each country. Mutations listed in order of frequency. Full data provided in Supplementary Table 6.



Supplementary Table 8 GDP Per Capita Verses Median MRSCA date for INH, RIF, SLIs, and FLQ.

Countries	GDP per capita	Median MRSCA Date INH (Years Ago)	Median MRSCA Date RIF (Years Ago)	Median MRSCA Date SLIs (Years Ago)	Median MRSCA Date FLQ (Years Ago)
China	10150	16.18	16.11	9.88	12.8
Peru	7140	19.04	6.52	2.61	2.39
Russia	11190	9.26	5.71	3.02	12.22
South Africa	6330	5.58	4.63	4.01	2.06
United Kingdom	42310	14.76	13.02	10.61	8.36
Uzbekistan	1480			2.45	
Romania	12510			11.74	10.19
Mali	933.88	4.97	3.4		
Netherlands	53020	11.59			
Moldova	3400		3.83		
Malawi	366.53	9.72	6.62		
South Korea	31940		17.66	16.08	11.95
Sierra Leone	516.76	6.85	5.24		
Germany	47790	12.48	13.86		
India	2200		11.27		





Supplementary Table 9 Commercial Diagnostics^{21,22,23,24,25}.

Drug	Gene	Location
Isoniazid	katG codon	315
Isoniazid	inhA promoter	-15,-16,-8
Rifamycin	rpoB codons	424-454
Aminoglycosides	rrs	1401,1402
Aminoglycosides	eis	-10 to -14, -37
Fluoroquinolones	gyrA	89-94
Fluoroquinolones	gyrB	500-541

*E-coli 505-534

Supplementary Table 10 Genes Searched Per Drug.

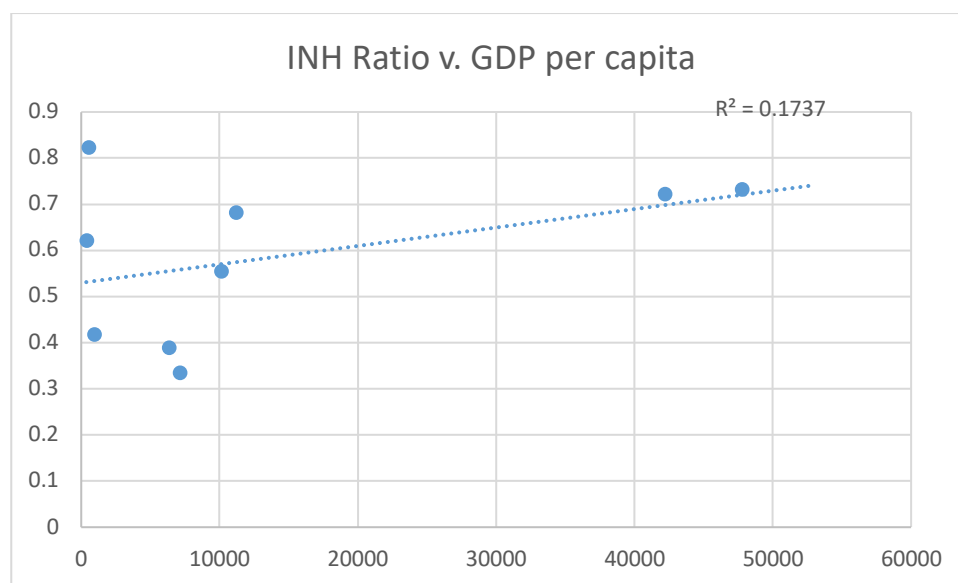
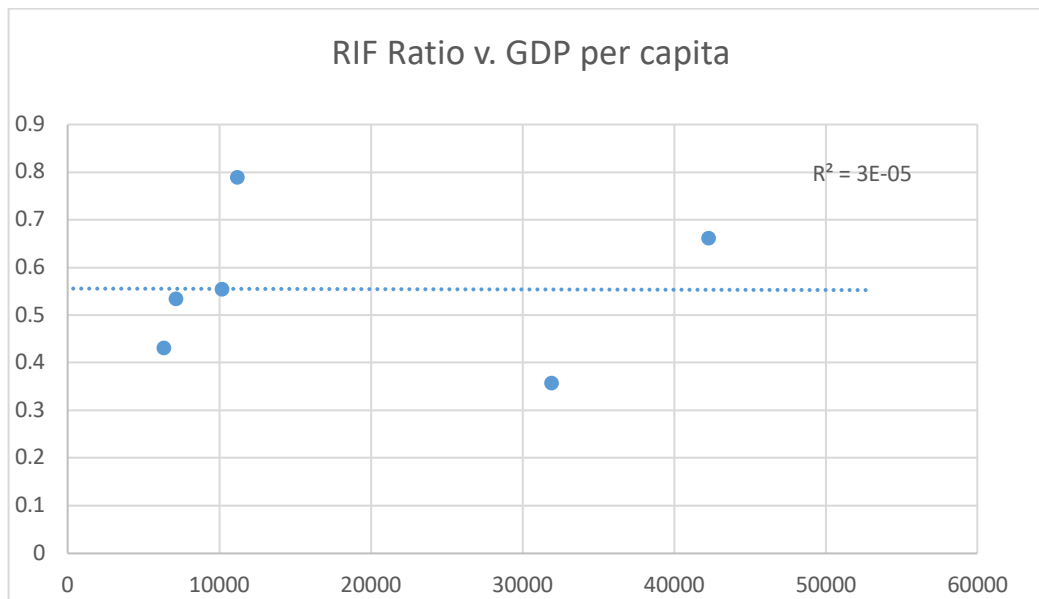
Genes searched for resistant mutations in “All WGS Test” (AMK: amikacin, PAS: para-aminosalicylic acid, EMB: ethambutol, CAP: capreomycin, KAN: kanamycin, CIP: ciprofloxacin, INH: isoniazid, STR: streptomycin, RIF: rifampicin, LEVO: levofloxacin, ETH: ethionamide, OFLX: ofloxacin, PZA: pyrazinamide)

Drug	Genes Searched
AMK	rrs
PAS	thyA, inter-thyA-Rv2765, folC, inter-thyX-hsdS.1
EMB	embA, embB, embC, iniB, inter-embC-embA
CAP	rrs, tlyA
KAN	rrs, inter-eis-Rv2417c

CIP	gyrB, gyrA
INH	inhA, iniB, embB, inter-Rv1482c-fabG1, ahpC, inter-oxyR'-ahpC, inter-embC-embA, kasA, katG, fabG1
STR	gid, rpsL, rrs, inter-rrs-rrl
RIF	rpoB
LEVO	gyrB, gyrA
ETH	inhA, inter-Rv1482c-fabG1, ethA, inter-ethA-ethR
OFLX	gyrB, gyrA
PZA	inter-pncA-Rv2044c, pncA, rpsA

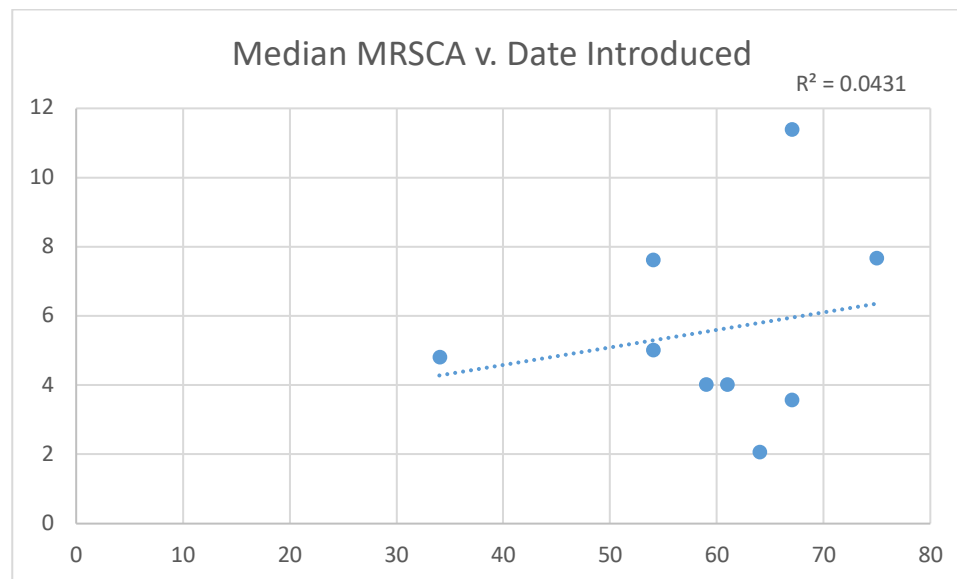
Supplementary Table 11 Transmission rate estimates and association with GDP by country. Transmission rate X/Y indicates X unique MRSCA dates/Y number of MRSCA dates for a specific country and drug. Legend: Unique = number of unique MRSCA dates for drug and country, total = total number of dated resistant isolates for drug and country, Lower Bound (LB) = (total-unique)/total, Note: Data for country fewer than 10 resistant isolates per drug not shown.

	Pooled			INH			RIF			PZA			ELO			STR			SLIs			EMB		
	LB			LB			LB			LB			LB			LB			LB			LB		
Country*	unique	total	Transmission	unique	total	Transmission	unique	total	Transmission	unique	total	Transmission	unique	total	Transmission	unique	total	Transmission	unique	total	Transmission	unique	total	Transmission
China	296	441	33%	56	101	45%	56	101	45%				32	44	27%	48	71	32%	33	36	8%	52	58	10%
Germany	90	113	20%	30	41	27%										40	50	20%						
India	13	20	35%																					
Malawi	44	68	35%	23	37	38%										15	23	35%						
Mali	18	30	40%	5	12	58%																		
Moldova	8	13	38%																					
Netherlands																								
Peru	1260	2116	40%	12	36	67%	382	715	47%	177	268	34%	59	68	13%	170	247	31%	194	251	23%	239	395	39%
Romania	29	41	29%													9	18	50%	13	15	13%			
Russia	302	394	23%	73	107	32%	56	71	21%	32	35	9%	19	19	0%	69	92	25%	10	12	17%	31	34	9%
Sierra Leone	48	64	25%	14	17	18%										16	27	41%						
South Africa	491	1015	52%	69	178	61%	72	167	57%	34	75	55%	79	128	38%	65	120	46%	86	192	55%	54	86	37%
South Korea	43	79	46%				5	14	64%										14	30	53%			
United Kingdom	299	416	28%	133	184	28%	49	74	34%	27	36	25%	27	37	27%	28	33	15%	14	17	18%	21	35	40%
Uzbekistan	31	36	14%																9	11	18%	14	17	18%
Total	2972	4846	39%	415	713	42%	620	1142	46%	270	414	35%	216	296	27%	460	681	32%	373	564	34%	411	625	34%



Supplementary Table 12 MRSCA age is not associated with the earliest date of drug introduction into clinical use.

Drug	Median MRSCA	Date Introduced Years ago
INH	11.4	67
RIF	7.61	54
PZA	3.56	67
FLQ	4.81	34
STR	7.67	75
SLIs	4.02	61
EMB	5.01	54
ETH	4.01	59
CYS	2.05	64



Supplementary Table 13 Geographic variance of resistance mutations

A: *katG* S315T mutation prevalence among INH resistant isolates in Russia versus Peru (Fisher P-value 1×10^{-12})

	Russia	Peru
Mutation	444	510
No Mutation	82	250

B: *fabG/inhA*-15 C>T mutation prevalence among INH resistant isolates in Russia versus Peru (Fisher P-value 7×10^{-9})

	Russia	Peru
Mutation	43	149
No Mutation	483	611

C: 25 mutations that varied geographically to a larger extent than the mutation *fabG/inhA* promoter -15C>T (SD 10%, Range 0%-39%) , their standard deviations, and ranges.

Mutation	Standard Deviation	Range of observed frequency across countries
INH_SNP_CN_2155168_CG_katG_S315T	12.49%	54.55%-93.75%
INH_SNP_CN_4247429_AG_embB_M306V	12.03%	0.0%-43.56%
INH_SNP_CN_4247431_GA_embB_M306I	11.88%	0.0%-42.63%
INH_SNP_CN_4247431_GC_embB_M306I	11.88%	0.0%-42.63%
PZA_SNP_CN_2289073_GC_pncA_H57D	15.07%	0.0%-46.67%
EMB_SNP_CN_4247429_AG_embB_M306V	12.48%	8.33%-53.37%
EMB_SNP_CN_4247431_GA_embB_M306I	11.95%	9.02%-49.58%
EMB_SNP_CN_4247431_GC_embB_M306I	11.95%	9.02%-49.58%

EMB_SNP_CN_4247431_GT_embB_M306I	11.95%	9.02%-49.58%
STR_SNP_CN_781687_AG_rpsL_K43R	21.03%	0.0%-81.85%
STR_SNP_N_1472359_A514C_rrs	12.00%	0.0%-50.4%
STR_SNP_N_1473246_A1401G_rrs	17.90%	0.0%-60.0%
STR_SNP_CN_4407927_TG_gid_E92D	31.17%	0.0%-87.95%
STR_SNP_N_1472362_C517T_rrs	11.32%	0.0%-38.89%
STR_SNP_CN_4407967_AG_gid_L79S	20.81%	0.0%-72.73%
ETH_SNP_P_1673425_CT.15_fabG1.inhA	14.48%	27.78%-62.5%
ETH_SNP_CN_4326333_CG_ethA_A381P	19.24%	0.0%-48.44%
ETH_SNP_CN_1674481_TG_inhA_S94A	11.89%	0.0%-31.46%
KAN_SNP_N_1473246_A1401G_rrs	14.83%	45.65%-85.71%
CAP_SNP_N_1473246_A1401G_rrs	19.62%	23.08%-85.84%
AMK_SNP_N_1473246_A1401G_rrs	20.11%	46.43%-100.0%
AMK_SNP_N_1472359_A514C_rrs	20.90%	0.0%-64.0%
CIP_SNP_CN_7582_AG_gyrA_D94G	26.57%	16.67%-78.26%
OFLX_SNP_CN_7582_AG_gyrA_D94G	12.60%	8.33%-50.0%
OFLX_SNP_CN_7570_CT_gyrA_A90V	11.83%	5.45%-46.15%

D: Six mutations that varied by lineage to a larger extent than the mutation *fabG/inhA* promoter -15C>T (SD 9.3%, Range 8.8%-33%), their standard deviations, and ranges.

Mutation	Standard Deviation	Range of observed frequency across lineages
INH_SNP_CN_2155168_CG_kat G_S315T	16.00%	40.28%-84.23%
INH_SNP_CN_2518919_GA_kas A_G269S	11.11%	0.0%-25.66%
PZA_SNP_CN_2288952_CT_pnc A_G97D	11.74%	0.0%-27.27%
EMB_SNP_CN_4247429_AG_e mbB_M306V	13.65%	18.87%-53.85%
STR_SNP_CN_781687_AG_rpsL _K43R	20.23%	4.76%-57.36%
STR_SNP_CN_4407927_TG_gid _E92D	43.23%	0.0%-99.83%

Supplementary Table 14 Sensitivity and Specificity of commercial tests for INH, RIF, SLIS, and FLQ in five countries with the largest number of phenotyped strains: Russia, South Africa, Peru, Uzbekistan, and United Kingdom

Legend: Sensitivity=Percent of resistant isolates classified as resistant, Specificity=Percent of susceptible isolates classified as susceptible

Drug	<u>commercial test-Peru</u>		<u>commercial test-Russia</u>		<u>commercial test-South Africa</u>		<u>commercial test-Uzbekistan</u>		<u>commercial test-United Kingdom</u>	
	Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity
INH	84% (641/760)	90% (151/168)	87% (459/526)	91% (281/310)	88% (168/190)	72% (444/619)	90% (238/264)	NA 0/0	87% (245/283)	99% (1559/1567)
RIF	90% (623/692)	89% (209/236)	82% (347/425)	88% (363/412)	95% (158/166)	70% (433/616)	95% (248/262)	50% (1/2)	91% (86/95)	99% (1742/1750)
FLQ	38% (46/121)	98% (788/801)	38% (51/133)	94% (275/292)	77% (86/111)	87% (510/585)	0% (0/7)	100% (213/213)	90% (26/29)	98% (295/301)
SLI	47% (101/214)	98%(692/708)	77% (63/82)	48%(163/343)	77% (101/131)	91%(486/536)	79% (44/56)	66% (109/164)	90% (9/10)	100%(83/83)

Supplementary References:

1. Schmieder R, Edwards R. Quality control and preprocessing of metagenomic datasets. *Bioinformatics*. 2011;27(6):863-864. doi:10.1093/bioinformatics/btr026
2. Wood DE, Salzberg SL. Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biology*. 2014;15(3):R46. doi:10.1186/gb-2014-15-3-r46
3. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv:13033997 [q-bio]*. March 2013. <http://arxiv.org/abs/1303.3997>. Accessed April 11, 2019.
4. *Picard Toolkit*. Broad Institute, GitHub repository: Broad Institute; 2019. <http://broadinstitute.github.io/picard/>.
5. Walker BJ, Abeel T, Shea T, et al. Pilon: An Integrated Tool for Comprehensive Microbial Variant Detection and Genome Assembly Improvement. *PLOS ONE*. 2014;9(11):e112963. doi:10.1371/journal.pone.0112963
6. Ezewudo M, Borens A, Chiner-Oms Á, et al. Integrating standardized whole genome sequence analysis with a global Mycobacterium tuberculosis antibiotic resistance knowledgebase. *Scientific Reports*. 2018;8(1):15382. doi:10.1038/s41598-018-33731-1
7. Coll F, McNerney R, Guerra-Assunção JA, et al. A robust SNP barcode for typing *Mycobacterium tuberculosis* complex strains. *Nature Communications*. 2014;5:4812. doi:10.1038/ncomms5812
8. Seabold S, Perktold J. Statsmodels: Econometric and Statistical Modeling with Python. 2010:5.
9. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*. 2014;30(9):1312-1313. doi:10.1093/bioinformatics/btu033
10. Yang Z. Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: Approximate methods. *J Mol Evol*. 1994;39(3):306-314. doi:10.1007/BF00160154
11. Suchard MA, Lemey P, Baele G, Ayres DL, Drummond AJ, Rambaut A. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evol*. 2018;4(1). doi:10.1093/ve/vey016
12. Dixit A, Freschi L, Vargas R, et al. Whole genome sequencing identifies bacterial factors affecting transmission of multidrug-resistant tuberculosis in a high-prevalence setting. *Scientific Reports*. 2019;9(1):5602. doi:10.1038/s41598-019-41967-8
13. Sukumaran J, Holder MT. DendroPy: a Python library for phylogenetic computing. *Bioinformatics*. 2010;26(12):1569-1571. doi:10.1093/bioinformatics/btq228
14. Oliphant T. *NumPy: A Guide to NumPy*. USA: Trelgol Publishing; 2006. <http://www.numpy.org/>.
15. Jones E, Oliphant T, Peterson P. SciPy: Open Source Scientific Tools for Python. January 2001.
16. World Economic Outlook, April 2019: Growth Slowdown, Precarious Recovery. IMF. <https://www.imf.org/en/Publications/WEO/Issues/2019/03/28/world-economic-outlook-april-2019>. Accessed April 11, 2019.

17. Kaufmann J, Schering AG. Analysis of Variance ANOVA. In: *Wiley StatsRef: Statistics Reference Online*. American Cancer Society; 2014. doi:10.1002/9781118445112.stat06938
18. Murray JF, Schraufnagel DE, Hopewell PC. Treatment of Tuberculosis. A Historical Perspective. *Annals ATS*. 2015;12(12):1749-1759. doi:10.1513/AnnalsATS.201509-632PS
19. Zhang Y, Shi W, Zhang W, Mitchison D. Mechanisms of Pyrazinamide Action and Resistance. *Microbiology Spectrum*. 2014;2(4). doi:10.1128/microbiolspec.MGM2-0023-2013
20. Alapi EM, Fischer J. Table of Selected Analogue Classes. In: *Analogue-Based Drug Discovery*. John Wiley & Sons, Ltd; 2006:441-552. doi:10.1002/3527608001.ch23
21. Rahman A, Sahrin M, Afrin S, et al. Comparison of Xpert MTB/RIF Assay and GenoType MTBDRplus DNA Probes for Detection of Mutations Associated with Rifampicin Resistance in Mycobacterium tuberculosis. *PLOS ONE*. 2016;11(4):e0152694. doi:10.1371/journal.pone.0152694
22. policy_statement.pdf. https://www.who.int/tb/features_archive/policy_statement.pdf. Accessed May 25, 2019.
23. Chen H-Y, Yu M-C, Huang W-L, et al. Molecular Detection of Rifabutin-Susceptible Mycobacterium tuberculosis. *Journal of Clinical Microbiology*. 2012;50(6):2085-2088. doi:10.1128/JCM.00652-12
24. Tagliani E, Cabibbe AM, Miotto P, et al. Diagnostic Performance of the New Version (v2.0) of GenoType MTBDRsl Assay for Detection of Resistance to Fluoroquinolones and Second-Line Injectable Drugs: a Multicenter Study. *Journal of Clinical Microbiology*. 2015;53(9):2961-2969. doi:10.1128/JCM.01257-15
25. Xie YL, Chakravorty S, Armstrong DT, et al. Evaluation of a Rapid Molecular Drug-Susceptibility Test for Tuberculosis. *New England Journal of Medicine*. 2017;377(11):1043-1054. doi:10.1056/NEJMoa1614915